

WO 99/32147

PCT/US98/27364

1

pVR 1012-GP(IC)

Sequence Listing ID No: 1

General Description

DNA pVR 1012-GP(IC)
Local object
Created: 09/14/98 04:17PM
Last Modification Date: ? (no data)
length: 7003 bp
storage type: Basic
form: Circular

Comments

Restriction Map

BglII: 1 site AGATCT
TCTAGA

Clal: 1 site ATCCAT
TAGCTA

Dralll: 1 site CACNNNGTG
GTGNNNCAC

EcoRV: 1 site GATATC
CTATAG

HindIII: 1 site AAGCTT
TTCGAA

HpaI: 1 site GTTAAC
CAATTG

KasI: 1 site GGCGCC
CCGCGG

KpnI: 1 site GGTACC
CCATGG

NarI: 1 site GGCGCC
CCGCGG

PmlI: 1 site CACGTG
GTGCAC

PstI: 1 site CTCGAG
GAGCTC

PvuI: 1 site CGATCG
GCTAGC

SacII: 1 site CCGCGG
GGCGCC

Sall: 1 site GTCGAC
CAGCTG

XmnI: 1 site GAANNNTTC
CTTNNNNAAG

EcoRI: 2 sites GAATTC
CTTAAG

NcoI: 2 sites CCATGG
GGTACC

NdeI: 2 sites CATATG
GTATAC

SphI: 2 sites GCATGC
CGTACG

XhoI: 2 sites CTCGAG
GAGCTC

BamHI: 3 sites CGATCC
CCTAGG

WO 99/32147

2

PCT/US98/27364

BclI: 3 sites TGATCA
 ACTAGT

Functional Map**CDS (4 signals)****CMV IE 5' UT**

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4020 End: 4572

Kan r

Start: 6068 End: 6690 (Complementary)

Misc_feature (2 signals)**CMV enhancer**

Start: 248 End: 885

GP(IC)

Start: 1870 End: 4019

Annotations

```

1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
   AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC
.....
51  GAGACGGTCA CAGCTTGTCT GTAAGCGSAT GCCGGGAGCA GACAAGCCCG
   CTCTGCCAGT GTCGAACAGA CATTCCGCTA CGGCCCTCGT CTGTTGCGGC
.....
101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
   AGTCCCGCGC AGTCGCCAC AACCGGCCAC AGCCCCGACC GAATTGATAC
.....

```

NdeI

```

151 CGGCATCAGA GCAGATTCTA CTGAGAGTCC ACCATATGCC GTGTGAAATA
   GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACGC CACACTTTAT
.....
201 CCGCACAGAT CCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
   GCGGTGTCTA CGCATTCTTC TTTTATGGCG TAGTCTAACC GATAACCGGT
.....
251 TTGCATACGT TGTATCCATA TCATAATATG TACATTATA TTGGCTCATG
   AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC
.....
301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
   AGGTTGTAAT GCGCGTACAA CTGTAACATA TAACTGATCA ATAATTATCA
.....
351 AATCAATTAC GGGGTCATTA GTTCAATAGC CATATATGCA GTTCCGCGTT
   TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA
.....
401 ACATAACTTA CCGTAAATGG CCGGCCTGGC TGACCGCCCA ACGACCCCGG
   TGTATTGAAT GCCATTACG GGGCGGACCG ACTGGCGGGT TGCTGGGGGC
.....
451 CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
   GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTCG GGTATCCCT
.....
501 CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
   GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTG ACGGGTGAAC
.....

```

NdeI

```

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
   CGTCACTGAG TTCACATAGT ATACGGTTCA TCGGGGGGAT AACTGCACTT
.....
601 TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
   ACTGCCATT ACCGGCGCGA CCGTAATACG GGTCAATGAT TGGAAATACCC
.....

```

NcoI

```

651 ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
   TGAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC
.....

```

NcoI

```

701 GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
   CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG
.....
751 ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
   TGCCCTATAA GGTTCAGAGG TGGGGTAAC TCAAGTTACCC TCAAACAAA
.....
801 GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCGCCCCCA
   CCGTGCTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT
.....

```

851 TTGACGCAAA TCGGCGGTAG GCGGTACCG TGGGAGGTCT ATATAAGCAG
AACITGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
TCGAGCAAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCGGGGAA
AACTGGAGGT ATCTTCTGTG GCCCTGCCTA GGTGGGAGGC CCCGGCCCTT

1001 CCGTGCAATTG GAACGCGGAT TCCCCGTCC AAGAGTGACG TAAGTACCGC
CCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCAATGCGC

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
GATATCTGAG ATATCCGTGT GGGGAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CTTATGCTA TAGGTGATGG
AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC

1151 CATAGCTTAG CCTATAGGTG TGGGTATTG ACCATTATTG ACCACTCCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGGAATAAC TGGTGAGGGG

1201 TATTGGTGAC CATACCTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCTTC AGAGACTGAC
GTTCATAGAG ATAACCGATA TACGTTATG AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
TGCCTGAGAC ATAAAAATGT CCTACCCAG GGTAAATAAT AAATGTTTAA

1351 CACATATACA ACAACGCGT CCCCCGTCC CGCAGTTT ATAAACATA
GTCTATATGT TGTTCGGCA GGGGACCGG GCGTCAAAA TAATTTGTAT

1401 GCGTGGGATC TCCACGCGAA TCTCGGTAC GTGTCCGGA CATGGGCTCT
CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGCCCT GTACCCGAGA

1451 TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
AGAGGCCATC GCCGCCCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
TCCCGGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTCCG CACAAGCCG
TGAATCCGTG TCGTGTACG GGTGTTGGT GTACACCGC GTGTTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCAG
ACCGCCATCC CATACACAGA CTTTACTCG CACCTCTAAC CCGAGCGTGC

1651 GGTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTCTTC TACGTCCGTC

1701 CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
GACTCAACAA CATACACTA TTCTAGTCT CCATTGAGGG CAACGCCACG

HpaI

1751 TGTTAACGGT GGAGCCCACT GTAGTCTGAG CAGTACTCGT TGCTGCCCGG
 ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACCACGGCGC

NcoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
 GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGCTGACA AGGAAAGGTA

SaiI

NcoI PstI PmlI BclI EcoRV

1851 GGGTCTTTTC TGCAGTCACC GTCGTGACA CGTGTGATCA GATATCGCGG
 CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

EcoRI

1901 CCGCGCGGGC GCTCTAGAAT TCTCTAATCA CAGTCATCAT GCGAGCGTCA
 GCGCGCGCGG CGAGATCTTA AGAGATTAGT GTCAGTAGTA CCCTCGCAGT

1951 GGGATTCTGC AATTGCCCGG TGAGCGCTTC AGGAAAACAT CTTTCTTTGT
 CCTTAAGACG TTAACGGGGC ACTCGCGAAG TCCTTTTGTA GAAAGAAACA

2001 TTGGGTAATA ATCCTATTCC ATAAAGTCTT TTCAATCCCG TTGGGGGTTG
 AACCCATTAT TAGGATAAGG TATTTACAAA AAGTAGGGC AACCCCAAC

2051 TACACAACAA TACCCTACAA GTGAGTGATA TTGACAAGTT TGTGTGCCGA
 ATGTGTTCTT ATGGGATGTT CACTCACTAT AACTGTTCAA ACACACGGCT

2101 GACAAACTCT CTTCAACTAG CCAATTGAAG TCAGTCGGGT TGAACCTGGA
 CTGTTTGAGA GAAGTTGATC GGTAACTTC AGTCAGCCCA ACTTGAACCT

2151 GGGCAATGGA GTAGCAACTG ATGTACCAAC GGCAACCAA AGATGGGGTT
 CCCGTTACCT CATCGTTGAC TACATGGTTG CCGTTGGTTT TCTACCCCAA

2201 TTCGAGCTGG TGTCCACCA AAGGTGGTAA ATTACGAAGC TGGAGAATGG
 AAGCTCGACC ACAAGGTGGT TTCCACCATT TAATGCTTCG ACCTCTTACC

2251 GCTGAGAACT GTTATAACCT GGCTATAAAG AAAGTTGATG GTAGTGAGTG
 CGACTCTTGA CAATATTGGA CCGATATTTC TTTCAACTAC CATCACTCAC

2301 CCTACCAGAA GCCCCTGAGG GAGTGAGGGA TTTTCCCCGT TGCCGCTATG
 GGATGGTCTT CGGGGACTCC CTCCTCCCT AAAAGGGGCA ACGGCGATAC

2351 TACACAAAGT CTCAGGAAGT GGACCATGCC CAGGAGGACT CGCCTTTCAC
 ATGTGTTTCA GAGTCCTTGA CCTGGTACGG GTCCTCCTGA CCGGAAAGTG

2401 AAAGAAGGAG CCTTCTTCCT GTATGACCGA CTCGCATCAA CAATCATTTA
 TTTCTTCCTC GGAAGAAGGA CATACTGGCT GAGCGTAGTT GTTAGTAAAT

2451 TCGGGGTACA ACCTTTGCCG AAGGAGTTAT TGCATTTCTG ATCTTGCCTA
 AGCCCCATGT TCGAAACGGC TTCCTCAATA ACGTAAAGAC TAGAACGGAT

2501 AGGCGCGAAA GGATTTTTTC CAGTCTCCTC CATTGCATGA GCCTGCCAAC
 TCCGCGCTTT CCTAAAAAG GTCAGAGGAG GTAACGTACT CGGACGGTTG

BamHI

2551 ATGACCACGG ATCCCTCCAG TTACTATCAC ACGACAACAA TAAACTACGT
 TACTCGTGCC TAGGGAGGTC AATGATAGTG TGCTGTTGTT ATTTGATGCA

 2601 GGTGATAAAT TTTGGAACCA ACACCACAGA GTTCTGTTC CAAGTCGATC
 CCAACTATTA AAACCTTGGT TGTGGTGTCT CAAAGACAAG GTTCAGCTAG

XhoI

2651 ATTTGACGTA TGTGCAGCTC GAGGCAAGAT TCACACCACA ATTCCTTGTC
 TAAACTGCAAT ACACGTCGAG CTCCGTTCTA AGTCTGGTGT TAAGGAACAG

 2701 CTCCTAAATG AAACCATCTA CTCTGATAAC CGCAGAAGTA ACACAACAGG
 GAGGATTTAC TTTGGTAGAT GAGACTATTG GCGTCTTCAT TGTGTTGTCC

 2751 AAAACTAATC TGGAAATAA ATCCCACTGT TCATACCAGC ATGGGTGAGT
 TTTTGATTAG ACCTTTTATT TAGGGTGACA ACTATGGTCC TACCCACTCA

 2801 GGCCTTTCTG GGAATAATA AAAACTTCAC AAAAACCTT TCAAGTGAAG
 CCCGAAAGAC CCTTTTATT TTTTGAAGTG TTTTGGGAA AGTTCAC TTC

 2851 AGTTGTCTTT CGTACCTGTA CCAGAAACCC AGAACCAGGT CCTTGACACG
 TCAACAGAAA GCATGGACAT GGTCTTTGGG TCTTGGTCCA GGAACGTGTC

 2901 ACAGCGACGG TCTCTCCTCC CATCTCCGCC CACAACCACG CAGGCGAAGA
 TGTCGCTGCC AGAGAGGAGG GTAGAGGCGG GTCTTGGTGC GTCCGCTTCT

 2951 CCACAAAGAA TTGGTTTCAG AGGATTCCAC TCCAGTGGTT CAGATGCAAA
 GGTGTTTCTT AACCAAAGTC TCCTAAGGTG AGGTCACCAA GTCTACGTTT

 3001 ACATCAAGSG AAAGGACACA ATGCCAACCA CAGTGACGGG TGTACCAACA
 TGTACTTCCC TTTCTGTGT TACGGTTGGT GTCACTGCC ACATGGTTGT

BclI

3051 ACCACACCCT CTCCATTTCC AATCAATGCT CGCAACACTG ATCATACCAA
 TGGTCTGGGA GAGGTAAAGG TTAGTTACGA GCGTTGTGAC TAGTATGGTT

 3101 ATCATTATC GGCCTGGAGG GGCCCCAAGA AGACCACAGC ACCACACAGC
 TAGTAAATAG CCGGACCTCC CCGGGGTTCT TCTGGTGTG TGGTGTGTCG

 3151 CTGCCAAGAC CACCAGCCAA CCAACCAACA GCACAGAATC GACGACACTA
 GACGGTTCTG GTGGTGGGT GGTGGTTGT CGTGTCTTAG CTGCTGTGAT

 3201 AACCCAACAT CAGAGCCCTC CAGTAGAGGC ACGGACCAT CCAGCCCCAC
 TTGGGTTGTA CTCTCGGGAG GTCATCTCCG TGCCCTGGTA GGTGCGGGTG

 3251 GGTCCCAAC ACCACAGAAA GCCACGCCGA ACTTGGCAAG ACAACCCAA
 CCAGGGGTTG TGGTGTCTTT CCGTGCGGCT TGAACCGTTC TGTGGGGTT

 3301 CCACACTCCC AGAACAGCAC ACTGCCGCCA GTGCCATTCC AAGAGCCGTG
 GGTGTGAGGG TCTTCTCTGT TGACGGCGGT CACGGTAAGG TTCTCGGCAC

 3351 CACCCCGACG AACTCAGTGG ACCTGGCTTC CTGACGAACA CAATACGGGG
 GTGGGGCTCC TTGAGTCACC TGGACCGAAG GACTGCTTGT GTTATGCCCC

BamHI

3401 GGTGACAAAT CTCCTGACAG GATCCAGAAG AAAGCGAAGG GATGTCACCTC
 CCACCTGTTTA GAGGACTCTC CTAGGTCTTC TTTCGCTTCC CTACAGTGAG

 3451 CCAATACACA ACCCAAATGC AACCCAAACC TGCCTATTG CACAGCCTTG
 GGTATATGTGT TGGGTTTACG TTGGGTTTGG ACGTGATAAC CTGTGCGAAC

 3501 GATGAGGGTG CTCCCATAGG TTTAGCCTGG ATACCATACT TCGGGCCAGC
 CTACTCCAC GACGGTATCC AAATCGGACC TATGGTATGA AGCCCGGTCG

 3551 AGCTGAGGGA ATTTACACTG AAGGCATAAT CGAGAATCAA AATGGATTGA
 TCGACTCCCT TAAATGTGAC TTCCGTATTA CCTCTTAGTT TTACCTAACT

 3601 TCTGTGGATT GAGGCAGCTG GCCAACGAAA CGACACAAGC TCTTCAATTG
 AGACACCTAA CTCCGTCGAC CGGTTGCTTT GCTGTGTTCC AGAAGTTAAC

 3651 TTCTTAAGGG CAACTACTGA GTTCCGTACA TTCTCTATAC TAAATCGGAA
 AAGAATTCCC GTTGATGACT CAACGCATGT AAGAGATATG ATTTAGCCTT

 3701 AGCAATAGAC TTCTTGCTCC AAAGATGGGG AGGAACATGT CACATTCTAG
 TCGTTATCTG AAGAACGAGG TTTCTACCCC TCCTTGATCA GTGTANGATC

 3751 GGCCTGATTG TTGCATTGAA CCCCAAGATT GGACCAAAAA TATCACTGAT
 CCGGACTAAC AACGTAACCTT GGGGTTCTAA CCTCGTTTTT ATAGTGACTA

BclI

3801 AAAATTGATC AAATAATCCA TGACTTTGTC GATAATAATC TTCCAAATCA
 TTTAACTAG TTTATTAGGT ACTGAAACAG CTATTATTAG AAGGTTTAGT

 3851 GAATGATGGC AGCAACTGGT GGAATGGATG GAAACAATGG GTTCCTGCTG
 CTTACTACCG TCGTTGACCA CCTGACCTAC CTTTGTTACC CAAGGACGAC

 3901 GAATAGGAAT CACAGGAGTA ATCAATGCTA TTATTGCTTT GCTGTGCATT
 CTTATCCTTA GTGTCTCAT TAGTAACGAT AATAACGAAA CGACACGTAA

EcoRI

3951 TGCAAATTCA TGCTTTGAAC TAATATAGCA TCATACTTTA GAATTCCTAGA
 ACGTTTAAGT ACGAACTTG ATTATATCGT AGTATGAAAT CTTAAGATCT

NarIKasIBamHI BglII

4001 CCAGGCGCCT GGATCCAGAT CTGCTGTGCC TTCTAGTTGC CAGCCATCTG
 GGTCCGCGGA CCTAGGTCTA GACGACACGG AAGATCAACG GTCGGTAGAC

 4051 TTGTTTGCCC CTCCCCGTG CTTTCTTGA CCCTGGAAGG TGCCACTCCC
 AACAAACGGG GAGGGGGCAC GGAAGGAAT GGGACCTTCC ACCGTGAGGG

 4101 ACTGTCCTTT CTAATAAAAA TGAGGAAATT GCATCGCATT CTCTGAGTAG
 TGACAGGAAA GGATTATTTT ACTCCTTTAA CGTAGCGTAA CAGACTCATC

 4151 GTGTCAATCT ATTCTGGGGG GTGGGGTGGG GCAGCACAGC AAGGGGGAGG
 CACAGTAAGA TAAGACCCCC CACCCCACCC CGTCGTGTCC TTCCCCCTCC

	<u>SphI</u>			<u>KpnI</u>	
4201	ATTGGAAGA	CAATAGCAGG	CATGCTGGGG	ATGCGGTGGG	CTCTATGGGT
	TAACCCCTTCT	GTTATCGTCC	GTACGACCCC	TACGCCACCC	GAGATACCCA
.....					
	<u>KpnI</u>				
4251	ACCCAGCTGC	TGAAGAATTG	ACCCGGTTTC	TCCTGGGCCA	GAAAGAAGCA
	TGGGTCCACG	ACTTCTTAAC	TGGGCCAAGG	AGGACCCCGT	CTTTCTTCGT
.....					
4301	CCCACATCCC	CTTCTCTGTG	ACACACCCCTG	TCCACGCCCC	TGGTCTTAG
	CCGTGTAGGG	GAAGAGACAC	TGTGTGGGAC	AGGTCCGGGG	ACCAAGAATC
.....					
4351	TCCAGCCCC	ACTCATAGGA	CATCATAGC	TCAGGAGCGC	TCCGCTTCA
	AAGGTCCGGC	TGAGTATCCT	GTGAGTATCG	AGTCCCTCCG	AGGCGGAAGT
.....					
4401	ATCCACCCC	CTAAAGTACT	TGGAGCGGTC	TCTCCCTCCC	TCATCAGCCC
	TAGGGTGGGC	GATTTTCATGA	ACCTCGCCAG	AGAGGGAGGG	AGTAGTCGGG
.....					
4451	ACCAAACCAA	ACCTAGCCTC	CAAGAGTGGG	AAGAAATTAA	AGCAAGATAG
	TGGTTTGGTT	TGGATCGGAG	GTTCTCACCC	TTCTTTAATT	TCGTCTATC
.....					
4501	GCTATTAAAT	GCAGAGCGAG	AGAAAATGCC	TCCAACATGT	CAGGAAGTAA
	CGATAATTCA	CGTCTCCCTC	TCTTTTACGG	AGGTTGTACA	CTCCTTCATT
.....					
	<u>XmnI</u>				
4551	TGAGAGAAAT	CATAGAATTT	CTTCCGCTTC	CTCGCTCACT	GACTCGCTGC
	ACTCTCTTTA	GTAATCTTAA	GAAGGCCAAG	GAGCGAGTGA	CTGAGCGACG
.....					
4601	CCTCGGTCGT	TGGGTGCGG	CGAGCGGTAT	CAGCTCACTC	AAAGGCGGTA
	CGAGCCAGCA	AGCCGACGCC	CCTCGCCATA	GTCGAGTGAG	TTTCCCCCAT
.....					
4651	ATACGTTAT	CCACAGAATC	AGGGGATAAC	GCAGGAAAGA	ACATGTGACC
	TATGCCAATA	GGTGTCTTAG	TCCCCTATTG	CGTCTTTCT	TGTACACTCG
.....					
4701	AAAAGGCCAG	CAAAAGGCCA	GGAACCGTAA	AAAGGCCGCG	TTGCTGGCGT
	TTTTCCGGTC	GTTTTCGGGT	CCTTGGCATT	TTTCCGGCGC	AACGACCGCA
.....					
4751	TTTTCCATAG	GCTCCGCCCC	CCTGACGAGC	ATCACAAAAA	TCGACGCTCA
	AAAAGGTATC	CGAGGCGGGG	GGACTGCTCG	TAGTGTTTT	AGCTCCGAGT
.....					
4801	AGTCAGAGGT	GCGGAAACCC	GACAGGACTA	TAAAGATACC	AGGCGTTTCC
	TCAGTCTCCA	CCGCTTTGGG	CTGTCTGAT	ATTTCTATGG	TCCGCAAAGG
.....					
4851	CCCTGGAAGC	TCCCTCGTGC	GCTCTCTGT	TCCGACCCCTG	CCGCTTACCG
	GGGACCTTCG	AGGGAGCAGC	CGAGAGGACA	AGGCTGGGAC	GGCGAATGGC
.....					
4901	GATACCTGTC	CCCCTTTCTC	CCTTCGGGAA	GCGTGGCGCT	TTCTCAATGC
	CTATGGACAG	GCGGAAAGAG	GGAAGCCCTT	CGCACCCGCA	AAGAGTTACC
.....					
4951	TCACGCTGTA	GGTATCTCAG	TTCGGTGTAG	GTCGTTGCT	CCAAGCTGGG
	AGTCCGACAT	CCATAGAGTC	AAGCCACATC	CAGCAAGCGA	GGTTCGACCC
.....					
5001	CTGTGTCAC	GAACCCCCCG	TTCAGCCCGA	CCGCTGCGCC	TTATCCGGTA
	GACACACGTG	CTTGGGGGGC	AAGTCGGGCT	GGCGACGCGG	AATAGGCCAT
.....					


```

5051 ACTATCGTCT TGAGTCCAAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA
    TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CGGTGACCGT
.....
5101 GCAGCCACTG GTAACAGGAT TAGCAGACCG AGGTATGTAG GCGGTGCTAC
    CGTCGGTGAC CATTTGCTTA ATCGTCTCGC TCCATACATC CGCCACCATG
.....
5151 AGAGTTCTTG AACTGGTGGC CTAACACGG CTACACTAGA AGGACAGTAT
    TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT TCCTGTCATA
.....
5201 TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGACTTGGT
    AACCATAGAC GCGACACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA
.....
5251 AGCTCTTGAT CCGGCAAAAC AACCACCGCT GGTAGCGGTG GTTTTTTTGT
    TCGAGAACTA GGCCTTTGT TTGGTGGCGA CCATCGCCAC CAAAAAACA
.....
5301 TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
    AACGTTGCTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTCTAGGAA
.....
5351 TGACTTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTAA
    ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT
.....
5401 GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT
    CCTTAAACC ACTACTCTAA TAGTTTTTCC TAGAAGTCCA TCTAGGAAA
.....
5451 AAATTAAAAA TGAAGTTTAA AATCAATCTA AAGTATATAT GAGTAACTT
    TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA
.....
5501 GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC
    CCAGACTGTC AATGGTTACG AATTAGTCAC TCCGTGSATA GAGTCGCTAG
.....
5551 TGCTATTTTC GTTCATCCAT AGTTGCCTGA CTCCGGGGGG GGGGGCGGCT
    ACAGATAAAG CAAGTAGSTA TCAACGGACT GAGGCCCCCC CCCCCGCGA
.....
5601 GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG ACTCATACCA GGCTGAATC
    CTCCAGACGG AGCACTTCTT CCACAACGAC TGAGTATGGT CCGGACTTAG
.....
5651 GCCCCATCAT CCAGCCAGAA AGTGAGGGAG CCACGGTTGA TGAGAGCTTT
    CGGGGTAGTA GGTGGGTCTT TCACTCCCTC GGTGCCAAT ACTCTGAAA
.....
5701 GTTGTAGGTG GACCAGTTGG TGATTTTGAA CTTTGTCTTT GCCACGGAAC
    CAACATCCAC CTGGTCAACC ACTAAAACCT GAAAACGAAA CCGTGCCTTG
.....
5751 GGTCTGCGTT GTCGGGAAGA TCGGTGATCT GATCCTTCAA CTCAGCAAAA
    CCAGACGCAA CAGCCCTTCT ACGCACTAGA CTAGGAAGTT GAGTCGTTTT
.....
5801 GTTCGATTTA TTCAACAAAG CCGCCGTCCC GTCAAGTCAG CGTAATGCTC
    CAACCTAAAT AACTGTGTTT GCGCGCAGG CAGTTCAGTC GCATTACGAG
.....
5851 TGCCAGTGTG ACAACCAATT AACCAATTCT CATTAGAAAA ACTCATCCAG
    ACGGTCACAA TGTGGTTAA TTGGTTAAGA CTAATCTTTT TGAGTAGCTC
.....
5901 CATCAATGA AACTGCAATT TATTCATATC AGGATTATCA ATACCATATT
    GTAGTTTACT TTGACGTTAA ATAAGTATAG TCCTAATAGT TATGCTATAA
.....
5951 TTTGAAAAAG CCGTTTCTGT AATGAAGGAG AAAACTCACC GAGGCAGTTC
    AAACTTTTTC GGCAGAGACA TACTTCCTC TTTTGAGTGG CTCCGTCAAG
.....

```

6001 CATAGGATGG CAAGATCCTG GTATCGGTCT GCGATTCCGA CTCGTCCAAC
GTATCCTACC GTTCTAGGAC CATAGCCAGA CCTAAGGCT GAGCAGGTTG

6051 ATCAATACAA CCTATTAAAT TCCCCTCGTC AAAAATAAGG TTATCAAGTC
TAGTTATGTT GGATAATTAA AGGGGAGCAG TTTTATTCC AATAGTTCAC

HindIII

6101 AGAAATCACC ATGAGTGACG ACTCAATCCG GTGAGAATCG CAAAAGCTTA
TCTTACTGG TACTCACTGC TGACTTAGGC CACTCTTACC GTTTTCGAAT

6151 TGCATTTCTT TCCAGACTTG TTCAACAGGC CAGCCATTAC GCTCGTCATC
ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG GTCCGTAATG CGAGCAGTAG

6201 AAAATCACTC GCATCAACCA AACCGTTATT CATTCGTGAT TGCGCCTGAG
TTTTAGTGAG CGTAGTTGGT TTGGCAATAA GTAAGCACTA ACCGGGACTC

PvuI

6251 CGAGACGAAA TACGCGATCG CTGTTAAAAG GACAATTACA AACAGGAATC
GCTCTGCTTT ATGCGCTAGC GACAACTTTC CTGTAAATGT TTGTCCTTAG

6301 GAATGCAACC GCGCGAGGAA CACTGCCAGC GCATCAACAA TATTTTCACC
CTTACGTTGG CCGCGTCCCT GTGACGCTCG CGTAGTTGTT ATAAAAGTGG

6351 TGAATCAGGA TATTCCTCTA ATACCTGCAA TGCTGTTTC CCGGGGATCG
ACTTAGTCCT ATAAGAAGAT TATGGACCTT ACGACAAAAG GGCCCCTAGC

6402 CAGTCGTGAG TACCATGCA TCATCAGGAG TACGGATAAA ATGCTTGATG
GTCACCACTC ATTGGTACGT AGTAGTCCTC ATGCCTATTT TACGAACCTAC

6451 GTCCGAAGAG GCATAAATTC CGTCAGCCAG TTTAGTCTGA CCATCTCATC
CAGCCTCTTC CGTATTTAAG GCAGTCGGTC AAATCAGACT GGTAGAGTAG

6501 TGTAACATCA TTGGCAACGC TACCTTTGCC ATGTTTCAGA AACAACTCTG
ACATTGTAGT AACCGTTGCG ATGGAACGG TACAAAGTCT TTGTTGAGAC

ClaI

6551 GCGCATCGGG CTTCCCATAC AATCGATAGA TTGTCCGACC TGATTGCCCG
CCCGTAGCCC GAAGGGTATG TTAGCTATCT AACAGCGTGG ACTAACGGGC

6601 ACATTATCGC GAGCCCATTT ATACCCATAT AAATCAGCAT CCATGTTGGA
TGTAATAGCG CTCGGGTAAA TATGGCTATA TTTAGTCGTA GGTACAACCT

XhoI

6651 ATTTAATCGC GGCCTCGAGC AAGACGTTTC CCGTTGAATA TGGCTCATAA
TABATTAGCG CCGGAGCTCG TTCTGCAAAG GGCAACTTAT ACCGAGTATT

6701 CACCCCTTGT ATTACTGTTT ATGTAAGCAG ACAGTTTTAT TGTTTCATGAT
GTGGGGAACA TAATGACAAA TACATTGCTC TGTCAAAATA ACAAGTACTA

DraIII

6751 GATATATTTT TATCTTGTGC AATGTAACAT CAGAGATTTT GAGACACAAC
CTATATAAAA ATAGAACACG TTACATTGTA GTCTCTAAAA CTCTGTGTTG

DraIII

6801 GTGGCTTTCC CCCCCCCCCC ATTATTGAAG CATTATCAG GGTATTGTC
CACCGAAAGC GGGGGGGGGG TAATAACTTC GTAAATAGTC CCAATAACAG
.....
6851 TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAATAGGG
AGTACTCGCC TATGTATAAA CTTACATAAA TCTTTTATT TGTATTATCC
.....
6901 GTTCCGCCCA CATTTCCTCCG AAAACTCCCA CCTGACGTCT AAGAAACCAT
CAAGGCGCGT GTAAAGGGGC TTTTCACGGT GGACTGCACA TTCTTTGGTA
.....
6951 TATTATCATG ACATTACCT ATAAAAATAG GCGTATCAG AGCCCTTTC
ATAATAGTAC TGTAAATTGA TATTTTATC CCCATAGTGC TCCGGGAAG
.....
7001 GTC
CAG
.....

pVR 1012-GP(S)

Sequence Listing ID No: 2

General Description

DNA pVR 1012-GP(S)
 Local object
 Created: 09/14/98 03:58PM
 Last Modification Date: ? (no data)
 length: 7073 bp
 storage type: Basic
 form: Circular

Comments

Restriction Map

Ball: 1 site TGGCCA
ACCGGT

BclI: 1 site TCATCA
ACTAGT

Clal: 1 site ATCGAT
TAGCTA

DraIII: 1 site CACNNNGTG
GTGNNNCAC

HindIII: 1 site AAGCTT
TTCGAA

KasI: 1 site GCGGCC
CCGCGG

KpnI: 1 site GGTACC
CCATGG

NarI: 1 site GCGGCC
CCGCGG

PmlI: 1 site CACGTG
GTGCAC

PvuI: 1 site CGATCG
GCTAGC

SacII: 1 site CCGCGG
GGCGCC

Sall: 1 site GTCGAC
CAGCTG

XbaI: 1 site TCTAGA
AGATCE

XmnI: 1 site GAANNNTTC
CTTNNNAAG

NdeI: 2 sites CATATG
GTATAC

EcoRV: 3 sites GATATC
CTAATG

SphI: 3 sites GCATGC
CGTACG

NcoI: 4 sites CCATGG
GGTACC

BamHI: 6 sites GGATCC
CCTAGG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4090 End: 4642

Kan r

Start: 6138 End: 6760 (Complementary)

Misc_feature (2 signals)

CMV enhancer

Start: 248 End: 885

GP(S)

Start: 1870 End: 4089

Annotations

1 TCCGCGCGTTT CCGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCC
 A3CGCGCCAAA GCCACTACTG CCACCTTTTG AGACTGTGTA CGTCGAGGGC

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCC
 CTCGCGCAST GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGC

101 TCAGGGCCCG TCAGCGGGTG TTGCGGGTG TCGGGGCTGG CTTAACTATG
 AGTCCCGCGC AGTCGCCCCAC AACCGCCCCAC AGCCCCGACC GAATTGATAC

NdeI

151 CGGCATCAGA GCAGATTGTA CTGAGACTGC ACCATATGCG GTGTGAAATA
 CCCGTAGTCT CGTCTAACAT GACTCTCAGC TGTATACGC CACACTTTAT

BalI

201 CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
 GCGGTGTCTA CGCATTCTTC TTTTATGGCG TAGTCTAAC GATAACCGGT

251 TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
 AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
 AGGTTGTAAT GCGGGTACAA CTGTAACATA TAACTGATCA ATAATTATCA

351 AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
 TTAGTTAATG CCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA

401 ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG
 TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC

451 CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
 GGGTAACGTC AGTTATTACT GCATACAAGG GTATCATGTC GGTATCCCT

501 CTTTCCATTG ACGTCAATGG GTGCAATAT TACGGTAAAC TGCCCACTTG
 GAAAGGTAAC TCCAGTTACC CACCTCATAA ATGCCATTG ACGGCTGAAC

NdeI

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
 CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT

601 TGACGGTAAA TGGCCCCGCT GGCATTATGC CCAGTACATG ACCTTATGGG
 ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGAATACCC

NcoI

651 ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
 TGAAAGCATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

NcoI

701 GTGATCGGGT TTTGGCAGTA CATCAATGGC CGTGGATACC GGTTCGACTC
 CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACGAG

751 ACGGGGATTT CCAAGTCTCC ACCCCATGA CGTCAATGGG AGTTTGTTTT
 TCCCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAACAAAA

801 GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT

851 TTGACGCAAA TCGGCGGTAG GCGTGACGG TGGGAGGTCT ATATAAGCAG
AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTCGGGTA GGTCCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT

1001 CCGTGCAATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCAATGGC

SphI

1051 CTATAGACTC TACAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
AAAACCGAAC CCCGGATATG TGGGGCCGAA GGAATACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTOCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

1201 TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTC CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CAACATCTC TATTGGCTAT ATGCCAATAC TCTGTCTTC AGAGACTGAC
GTTGATAGAG ATAACCGATA TACGCTTATG AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
TGCCCTGAGAC ATAAAAATG CCTACCCAG GGTAAATAAT AATGTTTAA

1351 CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
GTGTATATGT TGTTCGGCA GGGGGCACGG CGGTCAAAA TAATTTGTAT

1401 GCGTGGGATC TCCACCGGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
CGCACCCATG AGGTGCGCTT AGAGCCCATC CACAAGGCCT GTACCCGAGA

1451 TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCCTC
AGAGCCCATC GCCGCCCTCA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
TGAATCCGTG TCGTGTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCAG
ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGACGCTG

1651 GCTGACGAG ATGGAAGACT TAAGGCAGCG GCAGAAGAG ATCCAGGCAG
CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTTCTTC TACGTCCGTC

1701 CTGAGTTGTT GTATTCTGAT AACAGTCAGA GGTAACCTCC GTTGCCGTGC
GACTCAACAA CATAGACTA TTCTAGTCT CCATTGAGGG CAACGCCACG

1751 TGTAAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
ACAAATGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NeoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA

Sall

NeoI

PmlI BclI EcoRV

1851 GCGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACCTCAGTGG CAGCAGCTCT GCACACTAGT CTATAGCGCC

SplI

EcoRV

1901 CCGCTCTAGC TAGATGCATG CTCGAGCGGC CGCCAGTGTG ATGGATATCT
GGCGAGATCG ATCTACGTAC GAGCTCGCCG GCGGTCACAC TACCTATAGA

NeoI

1951 GCAGAACTCT ATCTTCAGGA TCTCGCCATG GAGGGTCTTA GCCTACTCCA
CGTCTTAAGA TACAAGTCCT AGAGCGGTAC CTCCCAGAA CTGATGAGGT

2001 ATTGCCCAGA GATAAACTTC GAAAAAGCTC TTTCTTTGTT TGGGTCTCA
TAACGGGTCT CTATTTAAAG CTTTTTCGAG AAAGAAACA ACCCAGTAGT

2051 TCTTATTICA AAAGGCCTTT TCCATGCCTT TGGGTGTTGT GACCAACAGC
AGAATAAAGT TTTCCGGAAG AGGTACGGAA ACCCAACAAC CTGGTTGTCTG

2101 ACTTTAGAAG TAACAGAGAT TGACCAGCTA GTCTGCAAGG ATCATCTTGC
TGAAATCTTC ATTGTCTCTA ACTGGTCTGAT CAGACGTTCC TAGTAGAACG

2151 ATCAACTGAC CAGCTGAAAT CAGTTGGTCT CAACCTCGAG GGGAGCGGAG
TAGTTGACTG GTCGACTTTA GTCAACCAGA GTTGAGCTC CCCTCGCCTC

EcoRV

2201 TATCTACTGA TATCCCATCT GCGACAAAGC GTTGGGGCTT CAGATCTGGT
ATAGATGACT ATAGGGTAGA CGCTGTTTCG CAACCCCGAA GTCTAGACCA

2251 GTGCCTCCCC AAGTGGTCAG CTATCAAGCA GGAGAATGGG CTGAAAATTG
CACGGAGGGG TTCACCAAGT GATACTTCGT CCTCTTACCC GACTTTTAC

2301 CTACAATCTT GAAATAAAGA AACCGGACGG GACCGAATGC TTACCCCCAC
GATGTTAGAA CTTTATTCTT TTGGCCTGCC CTCGCTTACG AATGGGGGTG

2351 CGCCGGATGG TGTCAGAGGC TTTCCAAGGT GCGGCTATCT TCACAAAGCC
GCGGCCTACC ACAGTCTCCG AAAGGTTCCA CGGCGATACA AGTGTTCGG

2401 CAAGGAACCG GGCCTGCCC GGGTGAATA GCCTTTCACA AGGATCGAGC
GTTCTTGGC CCGGGACGGG CCCACTGATA CGGAAAGTGT TCCTACCTCG

2451 TTTCTTCTC TATGACAGGC TGGCTTCAAC TGTAATTTAC AGAGGAGTCA
AAAGAAGGAG ATACTGTCCG ACCGAAGTTC ACATTAAATG TCTCTCAGT


```

2501 ATTTTCCTGA CCGGGTAATC GCA1TCTTGA TATTGGCTAA ACCAAAGGAA
TAAAACGACT CCCCCATTAG CGTAAGAACT ATAACCGATT TGGTTTCCTT
.....
2551 ACGTTCCCTC AATCACCACC CATTGAGAG GCAGCAAAC ACAC7GAAAA
TGCAAGGAAG TTAGTGGGGG GTAAGCTCTC CGTCGTTTGA TGTGACTTTT
.....
2601 TACATCAAGT TACTATGCCA CATCCTACTT GGAGTACGAA ATCGAAAATT
ATGTAGTTCA ATGATACGGT GTAGCATGAA CCTCATGCTT TAGCTTTTAA
.....
2651 TTGGTGCTCA ACAC7CCACG ACCCTTTTCA AAATTAACAA TAATAC7TTT
AACCACGAGT TGTGAGGTGC TGGGAAAAGT TTTAATTGTT ATTATGAAAA
.....
2701 GTTCTTCTGG ACAGGCCCCA CACGCCTCAG TTCCTTTTCC AGCTGAATGA
CAAGAAGACC TGTCCGGGGT GTGCGGAGTC AAGGAAAAGG TCGACTTACT
.....
2751 TACCATTCAA CTTACCAAC AGTTGAGCAA CACAAC7GGG AAAC7AATTT
ATCGTAAGTT GAAGTGCTTG TCAACTCGTT GTGTTGACCC TTTGATTAAA
.....
2801 GGACACTAGA TGCTAATATC AATGCTGATA TTGGTGAATG GGCTTTT7GG
CCTGTGATCT ACGATTATAG TTACGACTAT AACCAC7TAC CCGAAAAACC
.....
2851 GAAAATAAAA AAATCTCTCC GAACAAC7AC GTGGAGAAGA GCTGTCTTTC
CTTTTATTTT TTTAGAGAGG CTGTTGATG CACCTCTTCT CGACAGAAAG
.....
2901 GAAAC7TTAT CGCTCAACGA GACAGAAGAC GATGATGCCA CATCGTCGAG
CTT7GAAATA GCGAGTTGCT CTG7CTTCTG C7ACTACGCT GTAGCAGCTC
.....
2951 AACTACAAGG GGAAGAATCT CCGACCGGGC CACCAGGAAG TATTGCGACC
TTGATGTTTC CTTCTTTAGA GGCTGGCCCG G7GGTCC7TC A7AAGCC7GG
.....
3001 TGGTTCCAAA GGATTCCCTT GGGATGGTTT CATTGCACGT ACCAGAAGGG
ACCAAGGTTT CCTAAGGGGA CCCTACCAA GTAACGTGCA TGGTCTTCCC
.....
3051 GAAACAACAT TGCCGTCTCA GAATTCGACA GAAGGTCCAA GAGTAGATGT
CTTGT7TGA ACGGCAGAGT CTTAAGCTGT CTTCCAGCTT CTCATCTACA
.....
3101 GAATACTCAG GAAACTATCA CAGAGACAAC TGCAACAATC ATAGGCACTA
CTTATGAGTC CTTTGATAGT GTCTCTGTTG ACGTTGTTAG TATCCGTGAT
.....
3151 ACGGTAACAA CATGCAGATC TCCACCATCG GGACAGGACT GAGCTCCAGC
TGCCATTGTT GTACGTCTAG AGGTGCTAGC CCTGTCTGA CTCGAGGTCT
.....
                                NcoI
                                -----
3201 CAAATCCTGA GTTCCTCACC GACCATGGCA CCAAGCCCTG AACTCAGAC
GTTTAGGACT CAAGGAGTGG CTGCTACCGT GGTTCGGGAC TCTGAGTCTG
.....
3251 CTCCACAACC TACACACCAA AACTACCAGT GATGACCACC GAGGAACCAA
GAGCTGTTGG ATGTGTGCTT TTGATGGTCA C7ACTGGTGG CTCCTTGTTT
.....
3301 CAACACCACC GAGAACTCT CTGGCTCAA CAACAGAAGC ACCCACTCTC
GTG7GGTGG CTCTTTGAGA GGACCGAGTT GTCTCTTCG TGGGTGAGAG
.....
3351 ACCACCCAG AGAATATAAC AACACGGTT AAAACTGTTT GGGCACAAGA
TGGTGGGGTC TCTTATATTG TTGTCGCCAA TTTTGACAA CCCGTGTTCT
.....

```

3401 GTCCACAAGC AACGGTCTAA TAACTTCAAC AGTAACACGT ATTCTTGGCA
CAGGTGTTTC TTGCCAGATT ATTGAAGTTG TCATTGTCCA TAAGAACCCT

3451 GCCTTGGACT TCGAAAACGC AGCAGAAGAC AAGTTAACAC CAGGGCCACG
CGGAACCTGA AGCTTTTGCG TCGTCTTCTG TTCAATTGTG GTCCCGGTGC

3501 GGTAAATGCA ATCCCAACTT ACACTACTGG ACTGCACAAG AACAAACATAA
CCATTTACGT TAGGGTTGAA TGTGATGACC TGACGTGTTT TTGTTGTATT

BamHI

3551 TGCTGCTGGG ATTGCCTGGA TCCCGTACTT TGCACCGGGT GCAGAAGGCA
ACGACGACCC TAACGGACCT ACGGCATGAA ACCTGGCCCA CGTCTCCGT

3601 TATACACTGA AGGCCTTATG CACAACCAA ATGCCTTAGT CTGTGGACTC
ATATGTGACT TCCGGAATAC GTGTTGTTT TACGGAATCA GACACCTGAG

3651 AGACAACCTG CAAATGAAAC AACTCAAGCT CTGCAGCTTT TCTTAAGGGC
TCTGTTGAAC GTTTACTTTG TTGAGTTTCA GACGTCGAAA AGAATCCCG

3701 CACGACGGAG CTGCGGACAT ATACCATACT CAATAGGAAG GCCATAGATT
GTGCTGCCTC GACGCTGTA TATGGTATGA GTTATCCTC CGGTATCTAA

BamHI

3751 TCCTTCTGCG ACGATGGGGC GGGACATGTA GGATCCTGGG ACCAGATTGT
AGGAAGACCC TGCTACCCCG CCTGTACAT CCTAGGACCC TGGTCTAACA

3801 TGCATTGAGC CACATGATTG GACCAAAAAC ATCACTGATA AAATCAACCA
ACGTAACCTG GTGTACTAAC CTGGTTTTTG TAGTGACTAT TTTAGTTGGT

3851 AATCATCCAT GATTTCATCG ACAACCTTT ACCCAATCAG GATAATGATG
TTASTAGGTA CTAAGTAGC TGTGGGAAA TGGGTTAGTC CTATTACTAC

BamHI

3901 ATAATTGGTG GACGGGCTGG AGACAGTGA TCCCTGCAGG AATAGGCATT
TATTAACCAC CTGCCCAGAC TCTGTACCT AGGGACGTCC TTATCCGTAA

3951 ACTGGAATTA TTATTGCAAT CATTGCTCTT CTTTGCGTCT GCAAGCTGCT
TGACCTTAAT AATAACGTTA GTAACGAGAA GAAACGCAGA CGTTCGACGA

BamHI

4001 TTGTTGAATA TCAGAATTCC AGCACTGGCG GCCGTTACTA GTGGATCCGA
AACAACTTAT AGTCTTAAGG TCGTGACCGC CGGCAATGAT CACCTAGGCT

NarI

BamHI

XbaI

KasI

BamHI

4051 GCTCGGATCC AAGCTCTAGA CCAGCGGCCT GGATCCAGAT CTGCTGTGCC
CGAGCCTAGG TTGAGATCT GGTCCGCGGA CCTAGGTCTA GACGACACGG

4101 TTCTAGTTC CAGCCATCTG TTGTTTGGCC CTCCCCCGTC CCTTCCTTGA
AAGATCAACC GTCGGTAGAC AACAAACGGG GAGGGGGCAC GGAAGGAACT

4151 CCTTGAAGG TGCCACTCCC ACTGTCCTTT CCTAATAAAA TGAGGAAATT
GGGACCTTCC ACGGTGAGGG TGACAGGAAA GGATTATTTT ACTCCTTTAA

4201 GCATCGCATT GTCTGAGTAG GTGTCATTCT ATTCTGGGGG GTGGGGTGGG
CGTAGCGTAA CAGACTCATC CACAGTAAGA TAAGACCCCC CACCCACCCC

SphI

4251 CCAGCACAGC AAGCGGGAGG ATTGGGAAGA CAATAGCAGG CATGCTGGGG
CGTCGTGTCT TCCCCCTCC TAACCCTTCT GTTATCGTCC GTACGACCCC

KpnI

4301 ATGCGGTGGG CTCTATGGGT ACCCAGGTGC TGAAGAATTG ACCCGGTTC
TACGCCACCC GAGATACCCA TGGGTCCACG ACTTCTTAAC TGGGCCAAGG

4351 TCCTGGGCCA GAAAGAAGCA GGCACATCCC CTTCTCTGTG ACACACCCTG
AGGACCCGGT CTTTCTTCGT CCGTGTAGGG GAAGAGACAC TGTGTGGGAC

4401 TCCAGCCCC TGGTTCTTAG TTCCAGCCCC ACTCATAGGA CACTCATAGC
AGGTGCGGGG ACCAAGAATC AAGGTGCGGG TGAGTATCCT GTGAGTATCG

4451 TCAGGAGGCC TCCGCCCTCA ATCCACCCCG CTAAAGTACT TGGAGCGGTC
AGTCCCTCCC AGCGGGAAGT TAGGGTGGGC GATTTTCATGA ACCTCGCCAG

4501 TCTCCCTCCC TCATCAGCCC ACCAAACCAA ACCTAGCCTC CAAGAGTGGG
AGAGGGAGGG AGTAGTCGGG TGGTTGGTT TGGATCGGAG GTTCTCACCC

4551 AAGAAATTAA AGCAAGTAG GCTATTAAGT GCAGAGGGAG AGAAATGCC
TTCTTTAATT TCGTTCTATC CGATAATTCA CGTCTCCCTC TCTTTACGG

XmnI

4601 TCCAACATGT GAGGAAGTAA TCAGAGAAAT CATAGAATTT CTTCGGCTC
AGGTTGTACA CTCCTTCATT ACTCTCTTA GTATCTTAA GAAGGCCAAG

4651 CTCGCTCACT GACTCGCTGC GCTCGGTCTG TCGGCTCCGG CGAGCGGTAT
GAGCGAGTGA CTGAGCGAGG CGAGCCAGCA AGCCGACGCC CCTCGCCATA

4701 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC
GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG

4751 GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA
CGTCCTTTCT TGTACACTCG TTTTCCGGTC GTTTTCCGGT CCTTGGCATT

4801 AAAGGCCGCG TTCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC
TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGACTGCTCG

4851 ATCACAATAA TCGACGCTCA ACTCAGAGGT GGCGAAACCC GACAGGACTA
TAGTGTTTTT AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTGAT

4901 TAAAGATACC AGCGGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT
ATTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCAGC CGAGAGGACA

4951 TCCGACCCTG CCGCTTACCG GATACCTCTC CGCCTTCTC CTTTCGGGAA
AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT

5001 CCGTGGCGCT TTCTCAATCC TCACGCTGTA GGTATCTCAG TTCGGTGTAG
CGCACCGCCA AAGAGTTACG AGTGCGACAT CCATAGACTC AAGCCACATC

```

5051  GTCGTTCCGT  CCAAGCTGGG  CTGTGTGCAC  GAACCCCCCG  TTCAGCCCGA
      CAGCAAGCGA  GGTTCGACCC  GACACACGTG  CTTGGGGGGC  AAGTCGGGCT
.....
5101  CCGCTGCGCC  TTATCCGGTA  ACTATCGTCT  TGAGTCCAAC  CCGGTAAAGAC
      GCGGACGCGG  AATAGGCCAT  TGATAGCAGA  ACTCAGGTTG  GGCCATTCTG
.....
5151  ACGACTTATC  GCCACTGGCA  GCAGCCACTG  GTAACAGGAT  TAGCAGAGCG
      TGCTGAATAG  CCGTGACCGT  CGTCGGTGAC  CATGTCTCTA  ATCGTCTCGC
.....
5201  AGGTATGTAG  GCGGTGCTAC  AGAGTCTTTG  AAGTGGTGGC  CTAACTACGG
      TCCATACATC  CGCCACGATG  TCTCAAGAAC  TTCACCACCG  GATTGATGCC
.....
5251  CTACACTAGA  AGGACAGTAT  TTGGTATCTG  CGCTCTGCTG  AAGCCAGTTA
      GATCTGATCT  TCCTGTCTAT  AACCATAGAC  GCGAGACGAC  TTCGGTCAAT
.....
5301  CCTTCGGAAG  AAGAGTTGGT  AGCTCTTGAT  CCGGCAAAAC  AACCACCGCT
      GGAAGCCTTT  TTCTCAACCA  TCGAGAACTA  GGCCGTTTGT  TTGGTGGCGA
.....
5351  GGTAGCGGTG  GTTTTTTTGT  TTGCAAGCAG  CAGATTACGC  GCAGAAAAAA
      CCATCGCCAC  CAAAAAAACA  AACGTTCTGT  GTCTAATGCG  CGTCTTTTTT
.....
5401  AGGATCTCAA  GAAGATCCTT  TGATCTTTTC  TACGGGGTCT  GACGCTCAGT
      TCCTAGAGTT  CTTCTAGGAA  ACTAGAAAAG  ATGCCCCAGA  CTGCGAGTCA
.....
5451  GGAACGAAAA  CTCACGTTAA  GGGATTTTGG  TCATGAGATT  ATCAAAAAGG
      CCTTGCTTTT  GAGTGCAATT  CCCTAAAACC  AGTACTCTAA  TAGTTTTTCC
.....
5501  ATCTTCACCT  AGATCCTTTT  AAATTAAAAA  TGAAGTTTTA  AATCAATCTA
      TAGAAGTGGA  TCTAGGAAAA  TTTAATTTTT  ACTTCAAAAT  TTAGTCAAGT
.....
5551  AAGTATATAT  GAGTAAACTT  GGTCTGACAG  TTACCAATGC  TTAATCAGTG
      TTCATATATA  CTCATTTGAA  CCAGACTGTC  AATGGTTACG  AATTAGTCAC
.....
5601  AGGCACCTAT  CTCAGCGATC  TGTCTATTTT  GTTCATCCAT  AGTTCCTGA
      TCCGTGGATA  CAGTCGCTAG  ACAGATAAAG  CAAGTAGGTA  TCAACGGACT
.....
5651  CTCGGGGGGG  GGGGGGGCGT  GAGGTCTGCC  TCGTGAAGAA  GGTGTTGCTG
      GAGGCCCCCC  CCCCCCGCGA  CTCCAGACGG  AGCACTTCTT  CCACAACGAC
.....
5701  ACTCATACCA  GGCCTGAATC  GCCCCATCAT  CCAGCCAGAA  AGTGAGGGAG
      TGAGTATGGT  CCGGACTTAG  CGGGGTAGTA  GGTCGGTCTT  TCACTCCCTC
.....
5751  CCACGGTTGA  TGAGAGCTTT  GTTGTAGGTG  GACCAGTTGG  TGATTTTGAA
      GGTGCCAACT  ACTCTCGAAA  CAACATCCAC  CTGGTCAACC  ACTAAAACCT
.....
5801  CTTTGTCTTT  GCCACGGAAC  GGTCTGCGTT  GTCGGGAAGA  TCCCTGATCT
      GAAAACGAAA  CGGTGCCTTG  CCAGACGCAA  CAGCCCTTCT  ACGCACTAGA
.....
5851  GATCCTTCAA  CTCAGCAAAA  GTTCGATTTA  TTCAACAAAG  CCGCCGTCCC
      CTAGGAAGTT  GAGTCGTTTT  CAAGCTAAAT  AAGTTGTTTC  GCGGCGAGGG
.....
5901  GTCAAGTCAG  CGTAATGCTC  TGCCAGTGTT  ACAACCAATT  AACCAATTCT
      CAGTTCAGTC  GCATTACGAG  ACGGTCACAA  TGTGTTTAA  TTGTTAAGA
.....
5951  GATTAGAAAA  ACTCATCGAG  CATCAATGA  AACTGCAATT  TATTCATATC
      CTAATCTTTT  TGAGTAGCTC  GTAGTTTACT  TTGACGTTAA  ATAAGTATAG
.....

```

6001 AGGATTATCA ATACCATATT TTTGAAAAG CCGTTTCTGT AATGAAGGAG
TCCTAATAGT TATGGTATAA AAACCTTTTC GCCAAAGACA TTACTTCCTC

6051 AAAACTCACC GAGGCAGTTC CATAGCATGG CAAGATCCTG GTATCGGTCT
TTTTGAGTGG CTCGCTCAAG GTATCCTACC GTTCTAGGAC CATAGCCAGA

6101 GCGATTCCGA CTCGTCCAAC ATCAATACAA CCTATTAATT TCCCCTCGTC
CGCTAAGGCT GAGCAGGTTG TAGTTATGTT CGATAATTAA AGGGGACCAG

6151 AAAAATAAGG TTATCAAGTG AGAAATCACC ATGAGTGACG ACTGAATCCG
TTTTTATTCC AATAGTTCAC TCTTTAGTGG TACTCACTGC TGACTTAGGC

HindIII

6201 GTGAGAATGG CAAAAGCTTA TGCATTCTT TCCAGACTTG TTCAACAGGC
CACTCTTACC GTTTTCGAAT ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG

6251 CAGCCATTAC GCTCGTCATC AAAATCACTC GCATCAACCA AACCGTTATT
GTCCGTAATG CGAGCAGTAG TTTTAGTGAG CCTAGTTGGT TTGGCAATAA

PvuI

6301 CATTTCGTGAT TCGGCCTGAG CGAGACGAAA TACGCGATCG CTGTTAAAG
GTAAGCACTA ACGCGGACTC GCTCTGCTTT ATGCGCTAGC GACAATTTTC

6351 GACAATTACA AACAGGAATC GAATGCAACC GCGCCAGGAA CACTGCCAGC
CTGTTAATGT TTGTCTTAG CTACGTTGG CCGCGTCCTT GTGACGGTCG

6401 GCATCAACAA TATTTTCACC TGAATCAGGA TACTCTTCTA ATACCTGGAA
CGTAGTTGTT ATAAAGTGG ACTTAGTCCT ATAAGAAGAT TATGGACCTT

6451 TGCTGTTTTC CCGGGGATCG CAGTGGTGAG TAACCATGCA TCATCAGGAG
ACGACAAAAG GGCCCTAGC GTCACTCTC ATTGGTACGT AGTAGTCCTC

6501 TACGGATAAA ATGCTTGATG GTCGGAAGAG GCATAAATC CGTCAGCCAG
ATGCCTATT TACGACTAC CAGCCTTCTC CGTATTTAAG GCAGTCGGTC

6551 TTAGTCTGA CCATCTCATC TGTAAATCA TTGGCAACGC TACCTTTGCC
AAATCAGACT GGTAGAGTAG ACATTGTAGT AACCGTTGCG ATGGAAACGG

ClaI

6601 ATGTTTCAGA AACAACTCTG GCGCATCGGG CTTCCTATAC AATCGATAGA
TACAAAGTCT TTGTTGAGAC CGCGTAGCCC GAAGGGTATG TTAGCTATCT

6651 TTGTCCGACC TGATTGCCCG ACATTATCGC GAGCCCATTT ATACCCATAT
AACAGCGTGG ACTAACGGGC TGTAAATAGCG CTCGGGTAAA TATGGGTATA

6701 AATCAGCAT CCATGTTGGA ATTTAATCGC GGCCTCGAGC AAGACGTTTC
TTAGTCGTA GGTACAACCT TAAATTAGCG CCGGAGCTCG TTCTGCAAG

6751 CCGTTGAATA TGGCTCATAA CACCCCTTGT ATTACTGTTT ATGTAAGCAG
GGCAACTTAT ACCGAGTATT GTGGGGAACA TAATGACAAA TACATTCTGC

6801 ACAGTTTTAT TGTTTCATGAT GATATATTTT TATCTTGTGC AATGTAACAT
TGTCAAAATA ACAAGTACTA CTATATAAAA ATAGAACACG TTACATTGTA

DraIII

```
6851 CAGAGATTTT GAGACACAAC GTGGCTTTCC CCCCCCCCCC ATTATTGAAG
      CTCTCTAAAA CTCTGTGTTC CACCGAAAGG GGGGGGGGGC TAATAACTTC
.....
6901 CATTATCAG GGTATTGTC TCATGAGCGG ATACATATTT GAATGTATTT
      GTAAATAGTC CCAATAACAG AGTACTCGCC TATGTATRAA CTTACATAAA
.....
6951 AGAAAAATAA ACAAAATAGGG GTTCCCGCCA CATTCCCCG AAAAGTGCCA
      TCTTTTTATT TGTTTATCCC CAAGGCGCGT GTAAAGGGGC TTTTCACGCT
.....
7001 CCTGACGTCT AAGAAACCAT TATTATCATG ACATTACCT ATAAAAATAG
      GGAATGCAGA TTCTTTGGTA ATAATAGTAC TGTAATTGGA TATTTTATC
.....
7051 GCGTATCAG AGGCCCTTTC GTC
      CGCATAGTGC TCCGGGAAAG CAG
.....
```

pVR 1012-GP(Z)

General Description

DNA pVR 1012-GP(Z)
 Local object
 Created: 09/15/98 05:06PM
 Last Modification Date: ? (no data)
 length: 7285 bp
 storage type: Basic
 form: Circular

Comments

Sequence Listing ID No: 3

Restriction Map

DraIII: 1 site CACNANGTG
 GTGNNTCAC

HindIII: 1 site AAGCTT
 TTCGAA

HpaI: 1 site GTTAAC
 CAATTG

KasI: 1 site GGCGCC
 CCGCGG

NarI: 1 site GGCGCC
 CCGCGG

NotI: 1 site GCGGCCGC
 CGCCGCGC

PmlI: 1 site CACGTG
 GTGCAC

PvuI: 1 site CGATCG
 GCTAGC

SacII: 1 site CCGCGG
 GGCGCC

XbaI: 1 site TCTAGA
 AGATCT

XhoI: 1 site CTCGAG
 GAGCTC

EcoRV: 2 sites GATATC
 CTATAG

NcoI: 2 sites CCATGG
 GGTACC

NdeI: 2 sites CATATG
 GTATAC

SphI: 2 sites GCATGC
 CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1860

TbGH

Start: 4302 End: 4854

Kan r

Start: 6350 End: 6972 (Complementary)

Misc_f ature (2 signals)

CMV enhancer

Start: 248 End: 885

GP(Z)

Start: 1870 End: 4301

Annotations


```

1  TCGCGCGTTT CCGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCC
   AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC
.....
51  GAGACGGTCA CAGCTTGTCT GTAACCGGAT CCGGGGAGCA GACAAGCCCC
   CTCTGCCAGT GTCGAACAGA CATTTCGCTA CGGCCCTCGT CTGTTGCGGC
.....
101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
   AGTCCCGCGC AGTCGCCAC AACC GCCCAC AGCCCCGACC GAATTGATAC
.....

```

NdeI

```

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
   CCGTAGTCT CGTCTAACAT GACTCTCAG TGGTATACGC CACACTTTAT
.....
201 CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
   GGCGTGTCTA CGCATTCCCT TTTTATGGCG TAGTCTAACC GATAACCGGT
.....
251 TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
   AACGTATCCA ACATAGGTAT AGTATTATAC ATGTAATAT AACCAGATAC
.....
301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
   AGGTTGTAAT GCGGGTACAA CTGTAACATA TAACTGATCA ATAATTATCA
.....
351 AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
   TTAGTTAATG CCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA
.....
401 ACATAACTTA CCGTAATGCG CCCGCTGGC TGACCGCCCA ACGACCCCGG
   TGTATTGAAT CCCATTACCC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC
.....
451 CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
   GGGTAACGCG AGTTATTACT GCATACAAGG GTATCATTCG GGTTATCCCT
.....
501 CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
   GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTG ACGGGTGAAC
.....

```

NdeI

```

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
   CGTCATGTAG TTCACATAGT ATACGGTTCA TCGGGGGGAT AACTGCAGTT
.....
601 TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
   ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGAATATCCC
.....

```

NcoI

```

651 ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
   TGAAAGGATG AACCCTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC
.....

```

NcoI

```

701 GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
   CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACGTAG
.....
751 ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
   TGCCCCAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAAA
.....
801 GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
   CCGTGCTTTT AGTTGCCCTG AAAGGTTTAA CAGCATTGTT GAGGCGGGGT
.....

```

851 TTGACGCAAA TGGGCGGTAG GCGTGACGG TCGGAGGTCT ATATAAGCAG
 AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
 TCGAGCAAAT CACTTGCCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
 AACTGGAGGT ATCTTCTGTG CCCCTGGCTA GGTGGGAGGC GCCGGCCCTT

1001 CCGTGCAATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
 GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCAATGGC

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
 GACATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
 AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
 ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

1201 TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
 ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
 GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTACAA GGATGGGGTC CCATTTATTA TTTACAAATT
 TGCCTGACAC ATAAAAATGT CCTACCCAG GGTAAATAAT AAATGTTTAA

1351 CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTC ATTAAACATA
 GTGTATATGT TGTTGCGGCA GGGGGCACGG CCGTCAAAAA TAATTTGTAT

1401 GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
 CGCACCCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA

1451 TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
 AGAGGCCATC CCCGCCCTGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGCCAG
 TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAAATG CCRCCACCAC CAGTGTGCCG CACAAGGCCG
 TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGCAGATTG GGCTCGCACG
 ACCGCCATCC CATAACAGCA CTTTACTCG CACCTCTAAC CCGAGCGTGC

1651 GGTGACGCAG ATGGAAGACT TAAGCCAGCG GCAGAAGAAG ATGCAGGCAG
 CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTTCTTC TACGTCCGTC

1701 CTGAGTTGCT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
 GACTCAACAA CATAAGACTA TTCTCACTCT CCATTGAGGG CAACGCCACG

HpaI

1751 TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
ACAAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NcoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTCA TTGTCTGACA AGGAAAGGTA

NcoIPmlIEcoRVNotI

1851 GGGTCTTTTC TGCACTCACC GTCCTCGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACCTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

NarINotI XbaIKasI

1901 CCGCTCTAGA CCAGGCGCCT GGATCGATCC GCGATGAAGA TTAAGCCGAC
GGCGAGATCT GGTCCGCGGA CCTAGCTAGG CGCTACTTCT AATTCGGCTG

1951 AGTGAGCGTA ATCTTCATCT CTCTTAGATT ATTTGTTTTT CAGAGTAGGG
TCACTCGCAT TAGAAGTAGA GAGAATCTAA TAAACAAAAG GTCTCATCCC

2001 GTCGTCAGGT CCTTTTCAAT CGTGTAACCA AAATAAACTC CACTAGAAGG
CAGCAGTCCA GCAAAAGTTA GCACATTGGT TTTATTGAG GTGATCTTCC

2051 ATATTGTGGG GCAACAACAC AATGGGCGTT ACAGGAATAT TCCAGTTACC
TATAACACCC CGTTGTTGTG TTACCCGCAA TGTCCTTATA ACGTCAATGG

2101 TCGTGATCGA TTCAAGAGGA CATCATCTCT TCTTTGGGTA ATTATCCTTT
AGCACTAGCT AAGTCTCTCT GTAGTAAGAA AGAAACCCAT TAATAGGAAA

2151 TCCAAAGAAC ATTTTCCATC CCACTTGGAG TCATCCACAA TAGCACATTA
AGGTTTCTTG TAAAGGTAG GGTGAACCTC AGTAGGTGTT ATCGTGTAAT

2201 CAGGTTAGTG ATGTGCACAA ACTAOTTTGT CGTGACAAAC TGTCATCCAC
GTCCAATCAC TACAGCTGTT TGATCAAACA GCACTGTTTG ACAGTAGGTG

2251 AAATCAATTG AGATCAGTTG GACTGAATCT CGAAGGGAAT GGAGTGGCAA
TTTAGTTAAC TCTAGTCAAC CTGACTTAGA GCTTCCCTTA CCTCACCCTT

2301 CTGACGTGCC ATCTGCAACT AAAAGATGGG GCTTCAGGTC CGGTGTCCCA
GACTGCACGG TAGACGTTGA TTTTCTACCC CGAAGTCCAG GCCACAGGGT

2351 CCAAAGGTGG TCAATTATGA AGCTGGTGAA TGGGCTGAAA ACTGCTACAA
GGTTTCCACC AGTTAATACT TCGACCACTT ACCCGACTTT TGACGATGTT

2401 TCTTGAAATC AAAAAACCTG ACGGCAGTGA GTGTCTACCA GCAGCGCCAG
AGAACTTTAG TTTTGTGGAC TGCCCTCACT CACAGATGGT CGTCGCGGTC

2451 ACGGGATTCG GGGCTTCCCC CGGTGCCCGT ATGTGCACAA AGTATCAGGA
TGCCCTAACC CCCGAAGGGG GCCACGGCCA TACACGTGTT TCATAGTCCT

2501 ACGGGACCGT GTGCCGAGA CTTTGCCCTC CATAAAGAGG GTGCTTCTT
TGCCCTGGCA CACGGCTCTT GAAACGGAAG GTATTTCTCC CACGAAAGAA

2551 CCTGTATGAT CGACTTGCTT CCACAGTTAT CTACCGAGGA ACGACTTTCC
GGACATCTA GCTGAACGAA GGTGTCAATA GATGGCTCCT TGCTGAAAGC

2601 CTGAAGGTGT CGTTGCATTT CTGATACTGC CCCAAGCTAA GAAGGACTTC
GACTTCACAC GCAACGTAAA GACTATGACG GGGTTCGATT CTTCCTGAAG

2651 TTCAGCTCAC ACCCCTTGAG AGAGCCGGTC AATGCAACGG AGGACCCGTC
AAGTCGAGTG TGGCGAATC TCTCGGCCAG TTACGTTGCC TCCTGGGCAG

EcoRV

2701 TAGTGGCTAC TATTCTACCA CAATTAGATA TCAGGCTACC GGTTTTGGAA
ATCACCAGATG ATAAGATGGT GTTAATCTAT AGTCCGATGG CCAAAACCTT

2751 CCAATGAGAC AGAGTACTTG TTCGAGGTTG ACAATTTGAC CTACGTCCAA
GGTTACTCTG TCTCATGAAC AAGCTCCAAC TGTTAAACTG GATGCAGGTT

2801 CTTGAATCAA GATTACACAC ACAGTTTCTG CTCCAGCTGA ATGAGACAAT
GAACCTAGTT CTAAGTGTGG TGTCAAAGAC GAGGTCGACT TACTCTGTTA

2851 ATATACAAGT GGGAAAAGCA GCAATACCAC GGGAAAAC TAATTGGAAGG
TATATGTTCA CCCTTTTCCCT CGTTATGGTG CCCTTTTGAT TAAACCTTCC

2901 TCAACCCCGA AATTGATACA ACAATCGGGG AGTGGGCTT CTGGGAAACT
AGTTGGGGCT TTAACATATGT TGTTAGCCCC TCACCCGGAA GACCCTTTGA

2951 AAAAAAACC TCACTAGAAA AATTGCGAGT GAAGAGTTGT CTTTACAGT
TTTTTTTTGG AGTGATCTTT TTAAGCGTCA CTTCTCAACA GAAAGTGTC

3001 TGTATCAAAC GGAGCCAAAA ACATCAGTGG TCAGAGTCCG GCGCGAACTT
ACATAGTTTG CCTCGGTTTT TGTAGTCACC AGTCTCAGGC CGCGCTTGAA

3051 CTTCCGACCC AGGGACCAAC ACAACAAC TGAGACCACAA AATCATGGCT
GAAGGCTGGC TCCCTGGTTG TGTGTTGAC TTCTGGTGT TTAGTACCGA

3101 TCAGAAAATT CCTCTGCAAT GGTTCAGTG CACAGTCAAG GAAGCGAAGC
AGTCTTTTAA GGAGACGTTA CCAAGTTCAC GTGTCAGTTC CTTCCCTTCG

3151 TGCAGTGTCTG CATCTAACAA CCCTTGCCAC AATCTCCAGC AGTCCCAAT
ACGTCACAGC GTAGATTGTT GGAACGGTG TTAGAGGTGC TCAGGGGTTA

3201 CCCTCACAA CAAACCAGGT CCGGACAACA GCACCCATAA TACACCCGTG
GGGAGTGTTG GTTTGGTCCA GGCTGTGTGT CGTGGGTATT ATGTGGGCAC

3251 TATAAACTTG ACATCTCTGA GCCAACTCAA GTTGAACAAC ATCACCAGCAG
ATATTTGAAC TGTAAGAGCT CCGTTCAGTT CAACCTGTTG TAGTGGCGTC

3301 AACAGACAC GACAGCACAG CCTCCGACAC TCCCTCTGCC ACGACCCGAG
TTGTCTGTTG CTCTCGTGTC GGAGGCTGTG AGGGAGACGG TGCTGGCGTC

3351 CCGGACCCCC AAAAGCAGAG AACACCAACA CGAGCAAGAG CACTGACTTC
GGCCTGGGGG TTTTCGTCTC TTGTGGTTGT GCTCGTTCTC GTGACTGAAG

3401 CTGGACCCCG CCACCACAAC AAGTCCCCAA AACACAGCG AGACCGCTGG
GACCTGGGCG GGTGGTGTG TTCAGGGGTT TTGGTGTGCG TCTGGCGACC

```

3451 CAACAACAAC ACTCATCACC AAGATACCGG AGAAGAGAGT GCCAGCAGCG
    GTTG-TGTTG TGAGTAGTGG TTCTATGGCC TCTTCTCTCA CGGTCGTCGC
.....
3501 GGAAGCTAGG CTTAATTACC AATACTATTG CTGGAGTCGC ACGACTGATC
    CCTTCGATCC GAATTAATGG TTATGATAAC GACCTCAGCG TCCTGACTAG
.....
3551 ACAGGCGGGA GAAGAACTCG AAGAGAAGCA ATTGTCAATG CTCAACCCAA
    TGTCGCGCCT CTTCTTGAGC TTCTCTTCGT TAACAGTTAC GAGTTGGGTT
.....
3601 ATGCAACCCCT AATTTACATT ACTGGACTAC TCAGGATGAA GGTGCTGCAA
    TACGTTGGGA TTAATGTAA TGACCTGATG AGTCCTACTT CCACGACGTT
.....
3651 TCGGACTGGC CTGGATACCA TATTTCGGGC CAGCAGCCGA GGAATTTAC
    AGCCTGACCG GACCTATGGT ATAAAGCCCG CTCGTCGGCT CCCTTAAATG
.....
3701 ATAGAGGGCC TAATGCACAA TCAAGATGGT TTAATCTGTG GGTGAGACA
    TATCTCCCCG ATTACGTGTT AGTTCTACCA AATTAGACAC CCAACTCTGT
.....
3751 GCTGGCCAC CAGACGACTC AAGCTCTTCA ACTGTTCCCTG AGAGCCACAA
    CGACCGGTTG CTCTGCTGAG TTCGAGAAGT TGACAAGGAC TCTCGGTGTT
.....
3801 CTGAGCTACG CACCTTTTCA ATCCTCAACC GTAAGGCAAT TGATTCTTG
    GACTCGATGC GTGGAAAAGT TAGGAGTTGG CATTCCTTAA ACTAAAGAAC
.....
3851 CTGCAGCGAT GGGGCGGCAC ATGCCACATT CTGGGACCGG ACTGCTGTAT
    GACGTCGCTA CCCC GCCGTG TACGGTGTA GACCCTGGCC TGACGACATA
.....
3901 CGAACCACAT GATTGGACCA AGAACATAAC AGACAAAAT GATCAGATTA
    GCTTGGTGTA CTAACCTGGT TCTTGTAATTG TCTGTTTTAA CTAGTCTAAT
.....
3951 TTCATGATT TGTTGATAAA ACCCTTCCCG ACCAGGGGGA CAATGACAA
    AAGTACTAAA ACAACTATTT TGGGAAGGCC TGGTCCCCCT GTTACTGTTA
.....
4001 TGGTGGACAG GATGGAGACA ATGGATACCG GCAGGTATTG GAGTTACAGG
    ACCACCTGTC CTACCTCTGT TACCTATGGC CGTCCATAAC CTCAATGTCC
.....
4051 CGTTATAATT GCAGTTATCG CTTTATTCTG TATATGCAA TTTGTCTTT
    GCAATATTAA CGTCAATAGC GAAATAAGAC ATATACGTTT AAACAGAAAA
.....
4101 AGTTTTTCTT CAGATTGCTT CATGCAAAAG CTCAGCCTCA AATCAATGAA
    TCAAAAACAA GTCTAACGAA GTACCTTTTC GAGTCGGAGT TTAGTTACTT
.....
4151 ACCAGGATTT AATTATATGG ATTACTTGAA TCTAAGATTA CTGACAAAT
    TGGTCCTAAA TTAATATACC TAATGAAGTT AGATTCTAAT GAAGTCTTTA
.....
4201 GATAATATAA TACACTGGAG CTTTAAACAT AGCCAATGTG ATTCTAACTC
    CTATTATATT ATGTGACCTC GAAATTTGTA TCGGTTACAC TAAGATTGAG
.....
4251 CTTTAAACTC ACAGTTAATC ATAAACAAGG TTTGGTACCG AGCTCGAATT
    GAAATTTGAG TGTCATTAG TATTTGTTCC AAACCATGGC TCGAGCTTAA
.....
4301 ATCTGCTGTG CCTTCTAGTT GCCAGCCATC TGTGTTTGC CCCTCCCCCG
    TAGACGACAC GGAAGATCAA CGGTCGGTAG ACAACAAACC GGGAGGGGGC
.....
4351 TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA
    ACGGAAGGAA CTGGGACCTT CCACGGTGAG GGTGACAGGA AAGGATTATT
.....

```

4401 AATGAGGAAA TTGCATCCCA TTGTCTGAGT AGGTGTCAAT CTATTCTGGG
TTACTCCTTT AACGTAGCGT AACAGACTCA TCCACAGTAA GATAAGACCC

4451 GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATTGGGAA GACAATAGCA
CCCACCCAC CCCGTCGTGT CGTTCCTCCCT CCTAACCTT CTGTTATCGT

SphI

4501 GGCATGCTGG GGATCGGGTG GGCTCTATGG GTACCCAGGT GCTGAAGAAT
CCGTACGACC CCTACGCCAC CCGAGATACC CATGGGTCCA CGACTTCTTA

4551 TGACCCGGTT CCTCTGGGC CAGAAAGAAG CAGGCACATC CCCTTCTCTG
ACTGGGCCAA GGAGGACCCG GTCTTCTTC GTCCGTGTAG GGAAGAGAC

4601 TGACACACCC TCTCCAGCC CCTGGTTCTT AGTTCAGCC CCACTCATAG
ACTGTGTGGG ACABGTGCGG GGACCAAGAA TCAAGGTGCG GGTGAGTATC

4651 GACACTATA GCTCAGGAGG GCTCCGCTT CAATCCCACC CGCTAAAGTA
CTGTGAGTAT CGAGTCTCC CGAGCGGAA GTTAGGGTGG CCGATTTCAT

4701 CTTGGAGCGG TCTCTCCCTC CCTCATCAGC CCACCAAACC AAACCTAGCC
GAACCTCGCC AGAGAGGGAG CGAGTAGTCC GGTGGTTTGG TTTGGATCGG

4751 TCCAAGAGTG GGAAGAAATT AAAGCAAGAT AGGCTATTAA GTGCAGAGGG
AGGTTCTCAC CCTTCTTTAA TTCTGTTCTA TCCGATAATT CACGTCTCCC

4801 AGAGAAATG CCTCCAACAT GTGAGGAAGT AATGAGAGAA ATCATAGAAT
TCTCTTTTAC GGAGGTGTGA CACTCCTTCA TTAATCTCTT TAGTATCTTA

4851 TTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC
AAGAAGCCGA AGGAGCGAGT GACTGAGCGA CCGAGCCAG CAAGCCGACG

4901 GCGGAGCGGT ATCAGCTCAC TCAAGGCGG TAATACGGTT ATCCACAGAA
CCGCTCGCCA TAGTCGAGTG AGTTTCCGCC ATTATGCCAA TAGGTGTCTT

4951 TCAGGGGATA ACGCAGGAAA GAACATGTCA GCAAAACCCC AGCAAAAGGC
AGTCCCTAT TCGTCTCTT CTGTGACACT CGTTTTCCGG TCGTTTTCCG

5001 CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC
GTCCTTGGCA TTTTTCCGGC GCAACGACCG CAAAAAGGTA TCCGAGGCGG

5051 CCCCTGACGA GCATACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC
GGGACTGCT CGTAGTGTCT TTAGCTCGGA GTTCAGTCTC CACCGCTTG

5101 CCGACAGGAC TATAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT
GGCTGTCTTG ATATTTCTAT GGTCCGCAAA GGGGGACCTT CGAGGGAGCA

5151 GCGCTCTCCT GTTCCGACCC TCCCGCTTAC CGGATACCTG TCCGCCTTTC
CGCGAGAGGA CAAGGCTGGG ACGGCGAATG GCCTATGGAC AGCGGAAAG

5201 TCCCTTCGGG AAGCGTGCGG CTTTCTCAAT GCTCACGCTG TAGGTATCTC
AGGGAAGCCC TTCGCACCGC GAAAGAGTTA CGAGTGCGAC ATCCATAGAG

5251 AGTTCGGTGT AGGTCGTTCT CTCCAAGCTG GGCTGTGTGC ACGAACCCCC
TCAAGCCACA TCCAGCAAGC GAGGTTCCAC CCGACACACG TGCTTGGGGG

```

5301 CGTTCACCCC GACCGCTGCG CTTTATCCCG TAACTATCGT CTTGAGTCCA
    GCAAGTCGGG CTGGCGACGC GGAATAGGCC ATTGATAGCA GAACTCAGGT
.....
5351 ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG
    TGGGCCATTG TGTGCTGAAT AGCGGTGACC GTCGTGGGTG ACCATTGTCC
.....
5401 ATTAGCAGAG CGAGGTATGT AGCGGTGCT ACAGAGTTCT TGAAGTGGTG
    TAATCGTCTC GCTCCATACA TCCGCCACGA TGTCTCAAGA ACTTCACCAC
.....
5451 GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGGCCTCTGC
    CGGATTGATG CCGATGTGAT CTTCTGTCA TAAACCATAG ACCCGAGACG
.....
5501 TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA
    ACTTCGGTCA ATGGAAGCCT TTTCTCAAC CATCGAGAAC TAGGCCGTTT
.....
5551 CAAACCCACG CTGTAGACGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC
    GTTTGGTGGC GACCATCGCC ACCAAAAAAA CAAACGTTCC TCGTCTAATG
.....
5601 CCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTT TCTACGGGGT
    CGCGTCTTTT TTCTCTAGAG TTCTTCTAGG AAACCTAGAA AGATGCCCCA
.....
5651 CTGACCGTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCAATGAGA
    GACTGCGAGT CACCTTGCTT TTGAGTGCAA TTCCCTAAAA CCAGTACTCT
.....
5701 TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA AATGAAGTTT
    AATAGTTTTT CCTAGAAGTG GATCTAGGAA AATTTAATTT TTAATTCAAA
.....
5751 TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT
    ATTTAGTTAG ATTTATATA TACTCATTG AACCAAGACTG TCAATGGTTA
.....
5801 CCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCCTTCATCC
    CGAATTAGTC ACTCCGTGGA TAGAGTCGCT AGACAGATAA AGCAAGTAGG
.....
5851 ATAGTTGCCG GACTCCGCGG GGGGGGGGCG CTGAGGTCTG CCTCGTGAAG
    TATCAACGGA CTGAGGCCCG CCCCCCCCCG GACTCCAGAC GGAGCACTTC
.....
5901 AAGGTGTTGC TGACTCATA CAGGCCTGAA TCGCCCCATC ATCCAGCCAG
    TTCCACAACG ACTGAGTATG GTCCGGACTT AGCGGGGTAG TAGGTCGGTC
.....
5951 AAAGTGAGGG AGCCACGGTT GATGAGAGCT TTGTTGTAGG TGGACCAGTT
    TTTCACTCCC TCGGTCCCAA CTACTCTCGA AACAACATCC ACCTGGTCAA
.....
6001 GGTGATTTTG AACTTTTGCT TTGCCACGGA ACGGTCTGCG TTGTCGGGAA
    CCACTAAAC TTGAAAACGA AACGGTGCCCT TGCCAGACGC AACAGCCCTT
.....
6051 GATGCGTGAT CTGATCCTTC AACTCAGCAA AAGTTCGATT TATTCAACAA
    CTACGCACTA GACTAGGAAG TTGAGTCGTT TTCAAGCTAA ATAAGTTGTT
.....
6101 AGCCGCCGTC CCGTCAAGTC AGCGTAATGC TCTGCCAGTG TTACAACCAA
    TCGCGCGCAG GGCAGTTCAG TCGCATTACG AGACGGTCAC AATGTTGGTT
.....
6151 TTAACCAATT CTGATTAGAA AAACCTATCG AGCATCAAAT GAAACTGCAA
    AATGGTTAA GACTAATCTT TTTGAGTAGC TCGTAGTTTA CTTTGACGTT
.....
6201 TTTATTCATA TCAGGATTAT CAATACCATA TTTTGA AAAA AGCCGTTTCT
    AAATAAGTAT ACTCCTAATA GTTATGGTAT AAAA AACTTTT TCGGCAAGA
.....

```

6251 GTAATGAAGG AGAAAACCTCA CCGAGGCAGT TCCATAGGAT GGCAAGATCC
CATTACTTCC TCTTTTGAGT GGCTCCCTCA AGGTATCCTA CCGTTCCTAGG

6301 TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC AACCTATTAA
ACCATAGCCA GACGCTAAGG CTGACCAGGT TGTAAGTTATG TTGGATAATT

6351 TTTCCCCTCG TCAAAAATAA GGTATCAAG TGAGAAATCA CCATGAGTGA
AAAGGGGAGC AGTTTTTATT CCAATAGTTC ACTCTTTAGT GGTACTCACT

HindIII

6401 CGACTGAATC CGGTGAGAAT GGCAAAAGCT TATGCATTTC TTTCCAGACT
GCTGACTTAG GCCACTCTTA CCGTTTTCGA ATACGTAAAG AAAGGTCTGA

6451 TGTTC AACAG GCCAGCCATT ACGCTCGTCA TCAAAATCAC TCGCATCAAC
ACAAGTTGTC CGGTCGGTAA TGCAGCAGT AGTTTTAGTG AGCGTAGTTG

PvuI

6501 CAAACCGTTA TTCATTCTGT ATTGCCGCTG AGCGAGACGA AATACGCGAT
GTTTGCGCAAT AAGTAAGCAC TAACCGGGAC TCGCTCTGCT TTATGCCGTA

PvuI

6551 CGCTGTAAAG AGGACAATTA CAAACAGGAA TCGAATGCAA CCGGCGCAGG
CGGACAATTT TCCTGTTAAT GTTTGTCTCTT AGCTTACGTT GGCGCGCTCC

6601 AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG GATATTCTTC
TTGTGACGGT CGCGTAGTTG TTATAAAAGT GCACTTAGTC CTATAAGAAG

6651 TAATACCTGG AATGCTGTTT TCCCGGGGAT CGCACTGGTG AGTAACCATG
ATTATGGACC TTACGACAAA AGGGCCCTTA GCGTCACCAC TCATTGGTAC

6701 CATCATCAGG AGTACGGATA AAATGCTTGA TGGTCGGAAG AGGCATAAAT
GTAGTAGTCC TCATGCCCTAT TTTACGAACT ACCAGCCTTC TCCGTATTTA

6751 TCCGTCAGCC AGTTTACTCT GACCATCTCA TCTGTAACAT CATTGGCAAC
AGGCAGTCGG TCAAATCAGA CTGGTAGAGT AGACATTGTA GTAACCGTTG

6801 GCTACCTTTG CCATGTTTCA GAAACAACCTC TGGCGCATCG GGCTTCCCAT
CGATGGAAAC GGTACAAAGT CTTTGTGAG ACCGCGTAGC CCGAAGGGTA

6851 ACAATCGATA GATTCTCGCA CCTGATTGCC CGACATTATC CCGAGCCCAT
TGTTAGCTAT CTAACAGCGT GGACTAACGG CCGTAATAG CGCTCGGGTA

XhoI

6901 TTATACCCAT ATAAATCAGC ATCCATGTTG GAATTTAATC GCGGCCTCGA
AATATGGGTA TATTTAGTCG TAGGTACAAC CTTAAATTAG CGCCGGAGCT

XhoI

6951 GCAAGACGTT TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT
CGTTCTGCAA AGGGCAACTT ATACCGAGTA TTGTGGGGAA CATAATGACA

7001 TTATGTAACC AGACAGTTT ATTGTTCTAG ATGATATATT TTTATCTTGT
AATACATTCT TCTGTCAAAA TAACAAGTAC TACTATATAA AAATAGAACA

DraIII

```
7051  GCAATGTAAC ATCAGAGATT TTGAGACACA ACGTGGCTTT CCCCCCCCCC
      CGTTACATTG TAGTCTCTAA AACTCTGTGT TGCACCGAAA GGGGGGGGGG
.....
7101  CCATTATTGA AGCATTATC AGGGTTATTG TCTCATGAGC GGATACATAT
      GGTAAATACT TCGTAAATAG TCCCAATAAC AGAGTACTCG CCTATGTATA
.....
7151  TTGAATGTAT TTAGAAAAAT AAACAATAG GGGTTCCGCG CACATTTCCT
      AACTTACATA AATCTTTTAA TTTGTTTATC CCCAAGGCGC GTGTAAAGGG
.....
7201  CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAC
      GCTTTTCACG GTGGACTGCA GATTCTTTGG TAATAATAGT ACTGTAATTG
.....
7251  CTATAAAAAAT AGGCCTATCA CGAGGCCCTT TCGTC
      GATATTTTAA TCCGCATAGT GCTCCGGGAA AGCAG
.....
```

pVR 1012-SGP(Z)

General Description

DNA pVR 1012-SGP(Z)
 Local object
 Created: 09/14/98 04:29PM
 Last Modified: 09/15/98 04:50PM
 length: 7272 bp
 storage type: Basic
 form: Circular

Comments

Restriction Map

Dralll: 1 site CACNNNGTG
 GTGNNNCAC

Hindlll: 1 site AAGCTT
 TTCGAA

HpaI: 1 site GTTAAC
 CAATTG

KpnI: 1 site GGTACC
 CCATGG

NotI: 1 site GCGGCCGC
 CGCCGGCG

PmlI: 1 site CACGTG
 GTGCAC

PvuI: 1 site CGATCG
 GCTAGC

SacII: 1 site CCGCGG
 GGCGCC

XbaI: 1 site TCTAGA
 AGATCT

XhoI: 1 site CTCGAG
 GAGCTC

EcoRV: 2 sites GATATC
 CTATAG

NcoI: 2 sites CCATGG
 GGTACC

NdeI: 2 sites CATATG
 GTATAC

SphI: 2 sites GCATGC
 CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4289 End: 4841

Kan^r

Start: 6337 End: 6959 (Complementary)

Misc_feature (2 signals)

Sequence Listing ID No: 4

WO 99/32147

35

PCT/US98/27364

CMV enhancer

Start: 248 End: 885

SGP(Z)

Start: 1870 End: 4288

Annotations

```

1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
   AGCGCGCAAA GCCACTACTG CCACTTTTCG AGACTGTGTA CGTCGAGGGC
.....
51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
   CTCTGCCAGT GTCGAACAGA CATTCCGCTA CGGCCCTCGT CTGTTCCGGC
.....
101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
   AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAC
.....

```

NdeI

```

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
   GCCGTAGTCT CGTCTAACAT GACTCTCAGC TCGTATACGC CACACTTTAT
.....
201 CCSCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGCCCA
   GCGGTGTCTA CGCATTCTC TTTTATGGCG TAGTCTAACC GATAACCGGT
.....
251 TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
   AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC
.....
301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
   AGGTTGTAAT GCGCGTACAA CTGTAACTAA TAACTGATCA ATAATTATCA
.....
351 AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
   TTAGTTAATG CCCAGTAAT CAAGTATCGG GTATATACCT CAAGCGGCAA
.....
401 ACATAACTTA CCGTAAATGG CCCGCCGTCG TGACCGCCCA ACGACCCCG
   TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC
.....
451 CCCATTGACG TCAATAATGA CGTATGTTC CATAGTAACG CCAATAGGGA
   GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTGC GGTATATCCCT
.....
501 CTTCCTATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
   GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTC ACGGGTGAAC
.....

```

NdeI

```

551 GCAGTACATC AAGTGATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
   CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT
.....
601 TGACGGTAAA TGCCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
   ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGAATACCC
.....

```

NcoI

```

651 ACTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
   TCAAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC
.....

```

NcoI

```

701 GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTCGACTC
   CACTACGCCA AACCGTCAT GTAGTTACCC GCACCTATCG CCAACTGAG
.....
751 ACCGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
   TGCCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAA
.....
801 GGCACCAAAA TCAACGGGAC TTCCAAAAT GTCGTAACAA CTCCGCCCCA
   CCGTGTTTTT AGTTGCCCTG AAAGGTTTA CAGCATTTG GAGCGGGGT
.....

```

851 TTGACGCAAA TGGGGCGTAG GCCTGTACGG TGGGAGGTCT ATATAAGCAG
 AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTGCTC

 901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
 TCGAGCAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
 AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTGCGAGGC GCCGCCCCCT

 1001 CGGTGCATTG GAACGCGCAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
 GCCACGTAAC CTTGCCCTA AGGGGCACGG TTCTCACTGC ATTCAATGGC

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
 GATATCTGAG ATATCCGTGT GGGGAACCG AGAATACGTA CGATATGACA

 1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
 AAAACCGAAC CCCGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC

 1151 TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
 ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

 1201 TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
 ATAACCACTG CTATGAAAGG TAATGATTAG CTATTGTACC GAGAAACGGT

 1251 CAACATCTC TATTGGCTAT ATGCCAATAC TCTGTCTTC AGAGACTGAC
 GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG

 1301 ACGGACTCTG TATTTTACA GGATGGGGTC CCATTTAITA TTTACAATT
 TGCCTGAGAC ATAAAAATGT CCTACCCAG GGTAAATAAT AAATGTTAA

 1351 CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTMT ATTAAACATA
 CTGTATATGT TGTTCGGCA GGGGCACGG GCGTCAAAA TAATTTGTAT

 1401 GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
 CGCACCCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA

 1451 TCTCCGGTAG CGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
 AGAGGCCATC GCCGCCCTGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

 1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCTAACAG TGGAGGCCAG
 TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

 1551 ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
 TGAATCCGTG TCGTGTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC

 1601 TCGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
 ACCGCCATCC CATAACAGA CTTTACTCG CACCTCTAAC CCGAGCGTGC

 1651 GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
 CGACTGCGTC TACCTTCTGA ATTCCGTGCG CGTCTTCTTC TACGTCCGTC

 1701 CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTCC
 GACTCAACAA CATAAGACTA TTCTAGTCT CCATTGAGGG CAACGCCACG

HpaI

1751 TGTAAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCCGG
ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NcoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
CGCGGGTGGT CTGTATTATC GACTGTCTGA TTCTCTGACA AGGAAAGGTA

NcoIPmlIEcoRVNotI

1851 GGGCTTTTTC TGCAGTCACC GTCGTGACA CGTCTGATCA CATATCGCGG
CCCAGAAAAG ACCTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

NotI XbaI

1901 CCGCTCTAGA CCAGCGCCTT GGATCGAATT GATGAAGATT AAGCCGACAG
GGCTAGATCT GGTCCGCGGA CCTAGCTTAA CTACTTCTAA TTCGGCTGTC

1951 TGAGCGTAAT CTTTCATCTCT CTTAGATTAT TTGTTTCCA GAGTAGGGGT
ACTCGCATTA GAAGTAGAGA GAATCTAATA AACAAAAGGT CTCATCCCCA

2001 CSTCAGGTCC TTTTCAATCG TGTAACCAA ATAACTCCA CTAGAAGGAT
CCAGTCCAGG AAAAGTTAGC ACATGGGTTT TATTTGAGGT GATCTTCCTA

2051 ATTGTGGGCG AACACACAA TGGGCGTTAC AGGAATATTG CAGTTACCTC
TAACACCCCG TTGTGTGTT ACCCGCAATG TCCTTATAAC GTCAATGGAG

2101 GTGATCGATT CAAGAGGACA TCATTCTTTC TTGCGGTAAT TATCCTTTTC
CACTAGCTAA GTTCTCCTGT AGTACAAAAG AAACCCATTA ATAGGAAAAG

2151 CAAAGAACAT TTTCCATCCC ACTTGCAGTC ATCCACAATA GCACATTACA
GTTTCTTGTA AAAGGTAGGG TGAACCTCAG TAGGTGTTAT CGTGTAAATG

2201 GGTTAGTGAT GTCGACAAAC TAGTTTGTCTG TGACAAACTG TCATCCACAA
CCAATCACTA CAGCTGTTTG ATCAACAGC ACTGTTTGAC AGTAGGTGTT

2251 ATCAATTGAG ATCAGTTGGA CTGAATCTCG AAGGGAATGG AGTGGCAACT
TAGTTAATC TAGTCAACCT GACTTAGAGC TTCCCTTACC TCACCGTTGA

2301 GACGTGCCAT CTGCAACTAA AAGATGGGGC TTCAGGTCCG GTGTCCCACC
CTGCACGGTA GACGTTGATT TTCTACCCCG AAGTCCAGGC CACAGGGTGG

2351 AAGGTTGGTC AATTATGAAG CTGGTGAATG GGCTGAAAAC TCCTACAATC
TTTCCACCAG TTAATACTTC GACCACTTAC CCGACTTTTG ACGATGTTAG

2401 TTGAAATCAA AAAACCTGAC GGGAGTGAGT GTCTACCAGC AGCGCCAGAC
AACTTTAGTT TTTTGGACTG CCTCACTCA CAGATGGTCTG TCGCGGTCTG

2451 CGGATTCGGG GCTTCCCCCG GTGCCGGTAT GTGCACAAAG TATCAGGAAC
CCCTAAGCCC CGAAGGGGGC CACGGCCATA CACGTGTTTC ATAGTCCTTG

2501 GGCACCGTGT GCCGGAGACT TTCCCTTCCA TAAAGAGGGT GCTTTCTTCC
CCCTGGCACA CGCCCTCTGA AACGGAAGGT ATTTCTCCCA CGAAAGAAGG

2551 TGTATGATCG ACTTGCTTCC ACAGTTATCT ACCGAGGAAC GACTTTCGCT
ACATACTAGC TGAACGAAGG TGTCATAGA TGGCTCCTTG CTGAAAGCGA

2601 GAAGGTCTCG TTGCATTTCT GATACTGCCC CAAGCTAAGA AGGACTTCTT
CTTCCACAGC AACGTAAAGA CTATGACGGG GTTCGATTCT TCCTGAAGAA

2651 CAGCTCACAC CCCTTGAGAG AGCCGGTCAA TGCAACGGAG GACCCCTCTA
GTCCAGTGTG GGGAACTCTC TCGGCCAGTT ACCTTGCCCTC CTGGGCAGAT

EcoRV

2701 GTGGCTACTA TTCTACCACA ATTAGATATC AGGCTACCGG TTTTGGAACC
CACCGATGAT AAGATGGTGT TAATCTATAG TCCGATGGCC AAAACCTTGG

2751 AATGAGACAG AGTACTTCTT CGAGGTTGAC AATTTGACCT ACGTCCAAC
TTACTCTGTC TCATGAACAA GCTCCAACCTG TTAAACTGGA TGCAGGTTGA

2801 TGAATCAAGA TTCACACCAC AGTTTCTGCT CCAGCTGAAT GAGACAATAT
ACTTAGTTCT AACTGTGCTG TCAAAGACGA GGTCGACTTA CTCTGTATATA

2851 ATACAAGTGG GAAAAGGAGC AATACCACGG GAAAACAAAT TTGGAAGGTC
TATGTTTACC CTTTTCTCG TTATGGTGCC CTITTGATTA AACCTTCCAG

2901 AACCCCGAAA TTGATACAAC AATCGGGSAG TGGCCCTTCT GGGAAACTAA
TTGGGGCTTT AACTATGTTG TTAGCCCTC ACCCGGAAGA CCCTTTGATT

2951 AAAAACCTCA CTAGAAAAAT TCGCAGTGAA GAGTTGTCTT TCACAGTTGT
TTTTTGAGT GATCTTTTAA AGCGTCACTT CTCACAGAA AGTGTCAACA

3001 ATCAAACGGA GCCAAAAACA TCAGTGGTCA GAGTCCGGCG CGAACTTCTT
TAGTTTCCCT CGCTTTTGT AGTCACCAGT CTCAGGCCGC GCTTGAAGAA

3051 CCGACCCAGG GACCAACACA ACAACTGAAG ACCACAAAAT CATGGCTTCA
GGCTGGGTCC CTGTTGTGT TGTGACTTC TGGTGTTTA GTACCGAAGT

3101 GAAAAATTCCT CTGCAATGGT TCAAGTGCAC AGTCAAGGAA GGAAGCTGC
CTTTTAAGGA GACGTTACCA AGTTCACGTG TCAGTTCCTT CCCTTCGACG

3151 AGTGTCCAT CTAACAACCC TTGCCACAAT CTCCACGAGT CCCCAATCCC
TCACAGCGTA GATGTTGGG AACGGTGTAA GAGGTGCTCA GGGGTTAGGG

3201 TCACAACCAA ACCAGGTCCG GACAACAGCA CCCATAATAC ACCCGTGTAT
AGTGTGTTGTT TGGTCCAGGC CTGTTGTCTG GGGTATTATG TGGGCACATA

3251 AAACCTTGACA TCTCTGAGGC AACTCAAGTT GAACAACATC ACCGCAGAAC
TTTGAACGT AGAGACTCCG TTGAGTTCAA CTTGTTCTAG TGGCGTCTTG

3301 AGACAACGAC AGCACAGCCT CCGACACTCC CTCTGCCACG ACCGCAGCCG
TCTGTTGCTG TCGTGTCCGA GGCTGTGAGG GAGACGGTGC TGCCGTCCGC

3351 GACCCCCAAA AGCAGAGAAC ACCAACACGA CCAAGAGCAC TGACTTCCTG
CTGGGGGTTT TCGTCTCTTG TGGTTGTGCT CGTTCTCGTG ACTGAAGGAC

3401 GACCCCGCCA CCACAACAAG TCCCCAAAAC CACAGCGAGA CCGCTGGCAA
CTGGGGCGGT GGTGTTCTTC AGCGTTTTG GTGTCGCTCT GCGACCGTT

3451 CAACAACACT CATACCAAG ATACCGGAGA AGAGAGTGCC AGCAGCGGA
GTTGTTGTGA GTAGTGGTTC TATGGCCTCT TCTCTACCG TCGTCGCCCT

3501 AGCTAGGCTT AATTACCAAT ACTATTGCTG GAGTCGCAGG ACTGATCACA
TCGATCCGAA TTAATGGTTA TGATAACGAC CTCAGCGTCC TGACTAGTGT
.....
3551 GCGGGGAGAA GAACTCGAAG AGAAGCAATT GTCAATGCTC AACCCTAATG
CCGCCCTCTT CTTGAGCTTC TCTTCGTAA CAGTTACGAG TTGGGTTTAC
.....
3601 CAACCCTAAT TTACATTACT GGACTACTCA GGATGAAGGT GCTGCAATCG
CTTGGGATTA AATGTAATGA CCTGATGAGT CCTACTTCCA CGACGTTAGC
.....
3651 GACTGGCCTG GATACCATAT TTCGGGCCAG CAGCCGAGGG AATTACATA
CTGACCGGAC CTATGGTATA AAGCCCCGTC GTCCGGCTCC TTAATGTAT
.....
3701 GAGGGGCTAA TGCACAATCA AGATGGTTTA ATCTGTGGGT TGAGACAGCT
CTCCCCGATT ACGTGTACT TCTACCAAAT TAGACACCCA ACTCTGTGCA
.....
3751 GGCCAACGAG ACGACTCAAG CTCTCAACT GTTCCTGAGA GCCACAACCTG
CCGGTTGCTC TGCTGAGTTC GAGAAGTTGA CAAGGACTCT CGGTGTTGAC
.....
3801 ACCTACGCAC CTTTTCAATC CTCAACCGTA AGGCAATTGA TTTCTTGCTG
TCGATCGGTG GAAAGTTAG GAGTTGGCAT TCCGTAACT AAAGAACGAC
.....
3851 CAGCGATGGG GCGGCACATG CCACATTCTG GGACCGGACT GCTGTATCGA
GTCGCTACCC CGCCGTGTAC GGTGTAACAC CCTGGCCTGA CGACATAGCT
.....
3901 ACCACATGAT TGGACCAAGA ACATAACAGA CAAAATTGAT CAGATTATTC
TGGTGTACTA ACCTGGTCTT TGTATTGTCT GTTTAACTA GTCTAATAAG
.....
3951 ATGATTTTGT TGATAAAACC CTTCCGGACC AGGGGGACAA TGACAATTGG
TACTAAAACA ACTATTTTGG GAAGGCCTGG TCCCCCTGTT ACTGTAAACC
.....
4001 TGGACAGGAT CGAGACAATG GATACCGGCA GGTATTGGAG TTACAGGCGT
ACCTGTCTTA CCTCTGTAC CTATGGCCGT CCATAACCTC AATGTCCGCA
.....
4051 TATAATTGCA GTTATCGCTT TATTCTGTAT ATGCAAAATT GTCTTTTAGT
ATATTAACGT CAATAGCGAA ATAAGACATA TACGTTTAA CAGAAAATCA
.....
4101 TTTTCTTCAG ATTGCTTCAT GGAAAAGCTC AGCCTCAAAT CAATGAAACC
AAAAGAAGTC TAACGAAGTA CCTTTTCGAG TCGGAGTTA GTTACTTTGG
.....
4151 AGGATTTAAT TATATGGATT ACTTGAATCT AAGATTACTT GACAAATGAT
TCCTAAATTA ATATACCTAA TGAACCTAGA TTCTAATGAA CTGTTTACTA
.....
4201 AATATAATAC ACTGGAGCTT TAAACATAGC CAATGTGATT CTAACCTCTT
TTATATTATG TGACCTCGAA ATTGTATCG GTTACACTAA GATTGAGGAA
.....
4251 TAAACTCACA GTTAATCATA AACAAGGTTT GGAATTGATC TGCTGTGCCT
ATTGAGTGT CAATTAGTAT TTGTTCCAAA CCTTAACCTAG ACGACACGGA
.....
4301 TCTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCCGTGC CTTCCTTGAC
AGATCAACGG TCGGTAGACA ACAAACGGGG AGGGGGCACC GAAGGAACTG
.....
4351 CTTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATG
GGACCTTCCA CGGTGAGGGT GACAGGAAAG GATTATTTTA CTCCTTTAAC
.....
4401 CATCGCATTG TCTGAGTAGG TGTCAATCTA TTCTGGGGGG TGGGGTGGGG
GTACCGTAAC AGACTCATCC ACAGTAAGAT AAGACCCCCC ACCCCACCCC
.....

SphI

4451 CAGCACAGCA AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA
GTCGTGTCGT TCCCCCTCCT AACCCCTTCTG TTATCGTCCG TACGACCCCT

KpnI

4501 TCGGGTGGGC TCTATGGGTA CCCAGGTGCT GAAGAATTGA CCCGGTTCCT
ACGCCACCCG AGATACCCAT GGTCCACGA CTTCTTAAC TGGCCAAGGA

4551 CCTGGGCCAG AAAGAAGCAG GCACATCCCC TTCTCTGTGA CACACCCTGT
CGACCCGGTC TTTCTTCGTC CGTGTAGGGG AAGAGACACT GTGTGGGACA

4601 CCACGCCCTT GGTTCCTTAGT TCCAGCCCCA CTCATAGGAC ACTCATAGCT
GGTGGCGGCA CCAAGAATCA AGGTGGGGGT GAGTATCTG TGAGTATCGA

4651 CAGGAGGGCT CCGCCTTCAA TCCACCCCGC TAAAGTACTT GGAGCGGTCT
GTCCTCCCGA GCGGAAGTT AGGCTGGGCG ATTTTCATGAA CCTCGCCAGA

4701 CTCCCTCCCT CATCAGCCCA CCAACCCAAA CCTAGCCTCC AAGAGTGGGA
GAGGGAGGGA GTAGTCGGGT GGTTCGGTTT GGATCGGAGG TTCACACCT

4751 AGAAATTAA CCAAGATAGG CTATTAAGTC CAGAGGGAGA GAAATGCCT
TCATTAAATTT CGTCTATCC GATAATTCAC GTCTCCCTCT CTTTACGGA

4801 CCAACATGTG AGGAAGTAAT GAGAGAAATC ATAGAATTTC TTCCGCTTC
GGTTGTACAC TCCTTCATTA CTCTCTTAG TATCTTAAAG AAGGCGAAGG

4851 TCGCTCACTG ACTCGCTGCG CTCGCTCGTT CGGCTGCGGC GACCGGTATC
AGCGAGTGCAG TGAGCGACGC GAGCCAGCAA GCGGACGCCG CTCGCCATAG

4901 AGCTCACTCA AAGCGGGTAA TACGGTTATC CACAGAATCA GGGGATAACG
TCGAGTGAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT CCCCTATTGC

4951 CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAA
GTCCTTTCTT GTACACTCGT TTTCCGGTCG TTTTCCGGTC CTTGGCATT

5001 AAGGCCGCGT TGCTGGCGTT TTCCATAGG CTCCGCCCCC CTGACGAGCA
TTCCGGCGCA ACGACCCCAA AAAGGTATCC GAGGCGGGGG GACTGCTCGT

5051 TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
AGTGTTTTFA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC TGTCTGATA

5101 AAAGATACCA GCGCTTTCCC CTTGGAAGCT CCTCGTGCG CTCTCCTGTT
TTTCTATGGT CCCCRAAGGG GGACCTTCA GGGAGCACGC GAGAGGACAA

5151 CCGACCCCTG CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCCGGGAG
GGCTGGGACG CGGAATGGCC TATGGACAGG CGGAAAGAGG GAAGCCCTTC

5201 CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
CCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA AGCCACATCC

5251 TCGTTGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC
AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA AGTCGGGCTG

5301 CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCACCC CGGTAAGACA
CGGACCGGGA ATAGGCCATT GATACAGAA CTCAGGTGG GCCATTCTGT

```

5351  CGACTTATCG  CCACTGCCAG  CAGCCACTGG  TAACAGGATT  AGCAGAGCGA
      GCTGAATAGC  GGTGACCGTC  GTCGGTGACC  ATTGTCTCTA  TCGTCTCGCT
.....
5401  GGTATGTAGG  CCGTGCTACA  GAGTTCTTGA  AGTGGTGGCC  TAACTACGGC
      CCATACATCC  CCCACGATGT  CTCAAGAACT  TCACCACCGG  ATTGATGCCG
.....
5451  TACACTAGAA  GGACAGTATT  TGGTATCTGC  GCTCTGCTGA  AGCCAGTTAC
      ATGTGATCTT  CCTGTCATAA  ACCATAGACC  CGAGACCACT  TCGGTCAATG
.....
5501  CTTTCGGAAA  AGAGTTGGTA  CCTCTTGATC  CGGCAAACAA  ACCACCGCTG
      GAAGCCTTTT  TCTCAACCAT  CGAGAACTAG  CCCGTTTGTT  TGGTGGCGAC
.....
5551  GTAGCGGTGG  TTTTMTTGT  TGCAAGCAGC  AGATTACCGG  CAGAAAAAAA
      CATCGCCACC  AAAAAACAA  ACGTTCGTCC  TCTAATGCGC  GTCTTTTTTT
.....
5601  GGATCTCAAG  AAGATCCTTT  GATCTTTTCT  ACGGGGTCTG  ACGCTCAGTG
      CCTAGAGTTC  TTCTAGGAAA  CTAGAAAAGA  TGCCCCAGAC  TGCGAGTCAC
.....
5651  GAACGAAAAC  TCACGTTAAG  GGATTTTGGT  CATGAGATTA  TCAAAAAGGA
      CTTGCTTTTG  AGTGCAATTC  CCTAAAACCA  GTACTCTAAT  AGTTTTTCCT
.....
5701  TCTTCACCTA  GATCCTTTTA  AATTAAAAAT  GAAGTTTTAA  ATCAATCTAA
      AGAAGTGGAT  CTAGGAAAAT  TTAATTTTAA  CTTCAAAATT  TAGTTAGATT
.....
5751  AGTATATATG  AGTAAACTTG  GTCTGACAGT  TACCAATGCT  TAATCAGTGA
      TCATATATAC  TCATTTGAAC  CAGACTGTCA  ATGGTTACGA  ATTAGTCACT
.....
5801  GGCACCTATC  TCAGCGATCT  GTCTATTTCT  TTCATCCATA  GTTGCCTGAC
      CCGTGGATAG  AGTCGCTAGA  CAGATAAAGC  AAGTAGGTAT  CAACGGACTG
.....
5851  TCCCGGGGGG  GGGGGCGCTG  AGGTCTGCCT  CGTGAAGAAG  GTGTTGCTGA
      AGGCCCCCCC  CCCCCCGGAC  TCCAGACGGA  GCACTTCTTC  CACAACGACT
.....
5901  CTCATACCAG  GCCTGAATCG  CCCCATCATC  CAGCCAGAAA  GTGAGGGAGC
      GAGTATGGTC  CGGACTTAGC  GGGGTAGTAG  GTCGGTCTTT  CACTCCCTCG
.....
5951  CACGGTTGAT  GAGAGCTTTG  TTGTAGGTGG  ACCAGTTGGT  GATTTTGAAC
      GTGCCAATA  CTCTCGAAAC  AACATCCACC  TGGTCAACCA  CTAAACTTGT
.....
6001  TTTTGCTTTG  CCACGGAACG  GTCTGCGTTG  TCGGGAAGAT  GCGTGATCTG
      AAAACGAAAC  GGTGCCTTGC  CAGACGCAAC  AGCCCTTCTA  CGCACTAGAC
.....
6051  ATCCTTCAAC  TCAGCAAAAG  TTGGATTTAT  TCAACAAAGC  CGCCGTCCCG
      TAGGAAGTTG  AGTCGTTTTT  AAGCTAAATA  AGTTGTTTCG  CGGGCAGGGC
.....
6101  TCAAGTCAGC  GTAATGCTCT  GCCAGTGTTA  CAACCAATTA  ACCAATTCTG
      AGTTCAGTCG  CATTACGAGA  CGGTCACAAT  GTTGGTTAAT  TGTTTAAGAC
.....
6151  ATTAGAAAAA  CTCATCGAGC  ATCAAATGAA  ACTGCAATTT  ATTCATATCA
      TAATCTTTTT  GAGTAGCTCG  TAGTTTACTT  TGACCTTAAA  TAAGTATAGT
.....
6201  GGATTATCAA  TACCATATTT  TTGAAAAAGC  CGTTTCTGTA  ATGAAGCAGA
      CCTAATAGTT  ATGGTATAAA  AACTTTTTTC  GCAAAGACAT  TACTTCCTCT
.....
6251  AAACTCACCG  AGGCAGTTCC  ATAGGATGGC  AAGATCCTGG  TATCGGTCTG
      TTTGAGTGGC  TCCGTCAAGG  TATCCTACCG  TTCTAGGACC  ATAGCCAGAC
.....

```

6301 CGATTCCGAC TCGTCCAACA TCAATACAAC CTATTAATTT CCCCTCGTCA
GCTAAGGCTG AGCAGGTTGT AGTTATGTTG GATAATTAAA GGGGAGCAGT

6351 AAAATAAGGT TATCAAGTGA GAAATCACCA TGAGTGACGA CTGAATCCGG
TTTTATTCCA ATAGTTCACT CTTAGTGGT ACTCACTGCT GACTTAGGCC

HindIII

6401 TGAGAATGGC AAAAGCTTAT GCATTTCTTT CCAGACTTGT TCAACAGGCC
ACTCTTACCG TTTTCGAATA CGTAAAGAAA GGTCTGAACA AGTTGTCCGG

6451 AGCCATTACG CTCGTCATCA AAATCACTCG CATCAACCAA ACCGTTATTC
TCGGTAATGC GAGCAGTAGT TTTAGTGAGC GTAGTTGGTT TGGCAATAAG

PvuI

6501 ATTCTGTGATT GCGCCTGAGC GAGACGAAAT ACGCGATCGC TGTAAAAGG
TAAGCACTAA CGCGGACTCG CTCGTCTTTA TCGCTAGCG ACAATTTTC

6551 ACAATTACAA ACAGGAATCG AATGCAACCG GCGCAGGAAC ACTGCCAGCG
TGTTAATGTT TGCTCTTAGC TTACGTTGGC CGCGTCTTTG TGACGGTCCG

6601 CATCAACAAT ATTTTCACCT CAATCAGGAT ATTCTTCTAA TACCTGGAAT
GTAGTTGTTA TAAAAGTGGA CTTAGTCCTA TAAGAAGATT ATGGACCTTA

6651 GCTGTTTCC CGGGGATCGC AGTGGTGAGT AACCATGCAT CATCAGGAGT
CGACAAAAGC GCCCTAGCG TCACCACTCA TTGGTACGTA GTAGTCCTCA

6701 ACGGATAAAA TGCTTGATGG TCGGAAGAGG CATAAATTC GTCAGCCAGT
TGCCTATTTT ACGAACTACC AGCCTTCTCC GTATTTAAGG CAGTCGGTCA

6751 TTAGCTGAC CATCTCATCT GTAACATCAT TGGCAACGCT ACCTTTGCCA
AATCAGACTG GTAGAGTAGA CATTGTAGTA ACCGTTGCGA TGGAAACGGT

6801 TGTTTCAGAA ACAACTCTGG CGCATCGGGC TTCCCATACA ATCGATAGAT
ACAAAGTCTT TGTTGAGACC GCGTAGCCCG AAGGGTATGT TAGCTATCTA

6851 TGTCGCACCT GATTGCCCGA CATTATCGCG AGCCCATTTA TACCCATATA
ACAGCGTGGG CTAACGGGCT GTAATAGCGC TCGGGTAAT ATGGGTATAT

XhoI

6901 AATCAGCATC CATGTTGGAA TTTAATCGCG GCCTCGAGCA AGACGTTTCC
TTAGTCGTAG GTACAACCTT AATTAGCGC CGGAGCTCGT TCTGCAAAGG

6951 CGTTGAATAT GGCTCATAAC ACCCCTTGTA TTACTGTTTA TGTAAGCAGA
GCAACTTATA CCGAGTATTG TGGGGAACAT AATGACAAAT ACATTCTGCT

7001 CAGTTTTATT GTTCATGATG ATATATTTTT ATCTTGTCGA ATGTAACATC
GTCAAAATAA CAAGTACTAC TATATAAAAA TAGAACACGT TACATTGTAG

DraIII

7051 AGAGATTTTG AGACACAACG TGGCTTTCCC CCCCCCCCCA TTATTGAAGC
TCTCTAAAAC TCTGTGTTGC ACCGAAAGGG GGGGGGGGGT AATAACTTCG

7101 ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA
TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT

7151 CAAAAATAAA CAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTCCCAC
CTTTTTATTT GTTTATCCCC AAGGCGCGTG TAAAGGGGCT TTTCACGGTG

7201 CTGACGTCTA AGAAACCATT ATTATCATGA CATTACCTA TAAAAATAGG
GACTGCAGAT TCTTTGGTAA TAATAGTACT GTAATTGGAT ATTTTATCC

7251 CGTATCACGA GGCCCTTTCG TC
GCATAGTCCT CCGGGAAAGC AG